

NASA-CR-165910

NASA Contractor Report 165910

NASA-CR-165910
19830012174

A Study of Noise Metric and Tone
Correction Accuracy

LIBRARY COPY

MAY 26 1983

LANGLEY RESEARCH CENTER
LIBRARY, NASA
HAMPTON, VIRGINIA

B. M. Sullivan and J. E. Mabry

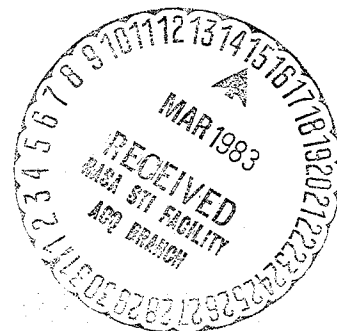
MAN-Acoustics and Noise, Incorporation
Seattle, Washington 98155

Contract NAS1-15955
April 1982



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665



NF01932

TABLE OF CONTENTS

List of Tables	ii
List of Figures.	iii
1.0 <u>INTRODUCTION.</u>	1
2.0 <u>LITERATURE SURVEY</u>	3
2.1 INTRODUCTION TO SURVEY	3
2.1.1 Loudness.	3
2.1.2 Noisiness	4
2.2 EXPERIMENTAL STUDIES	5
2.3 DIRECTION FOR THIS STUDY	26
3.0 <u>EXPERIMENT DESCRIPTION.</u>	38
3.1 SIGNALS.	38
3.2 EXPERIMENT DESIGN.	38
3.3 FACILITIES	40
3.4 PRESENTATION LEVELS.	40
3.5 INSTRUCTIONS	41
3.6 SUBJECTS	42
3.7 TRAINING	42
4.0 <u>DATA ANALYSIS</u>	43
4.1 NOISE METRIC DESCRIPTION	43
4.2 PSYCHOPHYSICAL METHOD.	47
5.0 <u>RESULTS</u>	51
6.0 <u>SUMMARY OF CONCLUSIONS.</u>	67

LIST OF TABLES

3-I.	EXPERIMENTAL SIGNALS.	39
4-I.	NOISE METRICS	44
5-I.	RANKING OF F-RATIOS FOR DATA COLLECTED AT NASA.	52
5-II.	RANKING OF F-RATIOS FOR DATA COLLECTED AT MAN	53
5-III.	RELATIVE ORDERING OF F-RATIOS FOR COMMON UNITS.	55
5-IV.	AVERAGE F-RATIOS FOR TWO SETS OF DATA	59
5-V.	TONE CORRECTION RANGES.	61

LIST OF FIGURES

2-1.	Difference Between 1/3 Octave Band Levels of (Tone plus Noise) and Noise Alone is 6.0 dB.	28
2-2.	Difference Between 1/3 Octave Band Levels of (Tone plus Noise) and Noise Alone is 15.0 dB.	29
2-3.	Difference Between 1/3 Octave Band Levels of (Tone plus Noise) and Noise Alone is 25.0 dB.	30
2-4.	Comparison of USASI Noise as Presented to Subjects at the NASA and MAN Facilities	32
4-1.	Derivation of Subjective dB.	48
5-1.	Results from NASA - Subjective dB for 32 Experimental Signals.	56
5-2.	Results from MAN - Subjective dB for 32 Experimental Signals.	57
5-3.	Signal 7 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations	63
5-4.	Signal 30 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations	64
5-5.	Signal 21 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations	65
5-6.	Signal 24 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations	66

A STUDY OF CALCULATION PROCEDURES FOR MEASURING ANNOYANCE RESPONSE TO AIRCRAFT FLYOVER SIGNALS

1.0 INTRODUCTION

There is an impressive number of methods (engineering calculation procedures) for quantifying annoyance response to individual aircraft flyover events. An element of many of these noise annoyance measurement approaches involves invoking a penalty for "tone" or discrete frequency characteristics in the flyover signal. The impetus for providing a penalty for "tone" or discrete frequency characteristics is, for the most part, due to human response studies which used artificial signals as stimulus materials. These early studies involving discrete frequency corrections were completed in the laboratory with pure tones, or narrow bands of noise superimposed on broad band artificial noise. Examples of studies utilizing pure tones are Little (1961, Ref. 1-1), Pearsons et al (1968, Ref. 1-2), and SAE (1972, Ref. 1-3). Other studies, including field studies where annoyance response to actual overflights was investigated and laboratory studies involving annoyance response to recordings of actual flyovers, often did not confirm the requirement for a tone correction. For a number of these studies, the tone correction to a particular engineering calculation procedure, such as PNdB, reduced the relationship between judged annoyance and the engineering measure of aircraft noise annoyance. Thus, the aim of this study is to reassess the requirement (including identification methods and quantification) for a discrete frequency correction to measurement methods for assessing noise annoyance response.

Two complementary methods are involved in developing a data base relative to the requirement for a "tone" or discrete frequency correction and the accuracy of various engineering calculation procedures. One approach uses a review and evaluation of previous experiments while the second approach involves an experiment consisting of two parts. The first part used recordings of actual flyover signals at a level that might be experienced in the open or out-of-doors area near an airport. The second part of the experiment

utilized the same signals modified to levels that would be experienced indoors near an airport.

REFERENCES

- 1-1. Little, J.W.: Human Response to Jet Engine Noise (Noise Control 7/3/11, May/June 1961).
- 1-2. Pearsons, K.S. et al: The Noisiness of Tones Plus Noise (NASA CR-1117, August 1968).
- 1-3. S.A.E. Research Project Committee R-6: Development of an Aircraft Flyover Noise Rating Scale - Report on Test No. 3: Constant Stimulus Difference Tone Study (Society of Automotive Engineers, Inc., New York, N.Y., April 1972).

2.0 LITERATURE SURVEY

2.1 INTRODUCTION TO SURVEY

2.1.1 Loudness

The quantification of human response to sound has had a long and complex history. Different parameters have been studied and different words have been used to describe these parameters, attempting to relate commonly perceived characteristics of sounds to scientifically measurable quantities.

One aspect that received early attention was "loudness," a subjective quality that could roughly be equated to the energy or overall sound pressure level (SPL) in a sound. However, the ear is differently sensitive to different frequencies, so that low-frequency sounds may sound "quieter" (less "loud") than higher frequency sounds of the same overall SPL. Fletcher and Munson (1933, Ref. 2-8) were early investigators of this phenomenon, and used a 1 KHz reference tone as the standard for comparison with tones of other frequencies. Many other investigators have worked in this field (See Karl Kryter's historical survey in "The Effects of Noise on Man," Academic Press, 1970.).

S.S. Stevens (1957, Ref. 2-57) published a procedure for combining bands of noise (narrow enough that the variation of loudness with frequency could be ignored) to evaluate the loudness of a broad band continuous-spectrum noise. The procedure is based on his "sone" scale, in which the unit of loudness is called a sone; one sone is ascribed to a 1 KHz tone set at an SPL of 40 dB, and a sound twice as loud as one sone is designated two sones, etc. The total loudness is given by:

$$\text{Loudness} = S_m + f (\Sigma S - S_m) \quad (\text{Equation 2-1})$$

where S_m is the sone value of the loudest band
 ΣS is the sum of the sones for all bands
 f is a fractional portion, dependent on
the bandwidth

Stevens later produced modifications in his method for calculating the sone value of a sound. Another investigator of loudness was E. Zwicker (1960, Ref. 2-62) whose more complex method of computing loudness takes into account the upward spread of masking, which affects perception.

The calculation of loudness using Steven's or Zwicker's method necessitates measuring band levels and performing a considerable amount of computation. To simplify this, and to enable an approximation of the subjective quantification of "loudness" to be measured with an electronic instrument, "weighting" responses were defined which accorded weighted intensity values to the frequency components in a sound to match the equal-loudness contours of Fletcher and Munson. Three responses were standardized, and electronic networks which evaluate sounds in accordance with these standards are built into most sound level meters for immediate evaluation of sounds. The three responses are the 'A'-weighting, which corresponds more or less to the 40 phon contour and therefore evaluates sounds in a manner similar to the way the human ear responds to quiet sounds, the 'B'-weighting, corresponding more or less to the 70 phon contour (representing medium-level sounds) and the 'C'-weighting, corresponding to the 100 phon contour, for evaluation of loud sounds. Sound levels measured with one of these weightings are expressed in dBA, dBB or dBC.

2.1.2 Noisiness

Other workers extended the concepts of human response to sound, to investigate the "unwantedness" of sound; in this context, the word "noise" is often used to imply "unwanted sound." People have noticed that a "loud" sound may be not unpleasing, whereas a "quiet" sound under certain circumstances may be very annoying. Obviously the information carried by the sound and the context in which it is heard will affect the hearer very strongly, but, even without these circumstances, for a sound carrying no particular information heard

under neutral conditions, there remained a residue of opinion that some sounds differed inherently in "noisiness" from others in a manner that did not simply equate with "loudness."

One of the most prominent investigators in this field who defined the concept of "perceived noisiness" is K. Kryter, who in 1959 (Ref. 2-14) proposed a scaling procedure similar to that of S.S. Stevens, using a similar formula to predict the total noisiness of a broad band noise from the noisiness (in "noys") of the individual bands. The equation analogous to equation 2-1 for 1/3 octave band levels is:

$$PNdB = N_{\max} + 0.15 (\Sigma N - N_{\max}) \quad (\text{Equation 2-2})$$

The noisiness in "noys" can be converted into a dB-type scale, giving the perceived noise level in PNdB; in the same way, the loudness of a sound in "sones" can be converted to a dB-scale of loudness, or "perceived level" (PL)[sometimes designated "perceived loudness level" (PLL) or "loudness level" (LL)] in phons (or PLdB), using methods based on Stevens' or Zwicker's procedures.

2.2 EXPERIMENTAL STUDIES

The importance of the word used to describe the quality of the sound that is to be judged, be it "loud," or "unwanted," "noisy," "disturbing," "objectionable," "unacceptable," "unpleasant," etc., was noted in respect to aircraft noise by Copeland et al (1960, Ref. 2-7) who presented to a large jury of listeners recordings of civil aircraft (including turboprop and turbofan noise), together with synthesized flyover noises of two kinds, one being a real jet-rig noise with artificial Doppler effect and rise and fall of level to simulate a flyover, and the other being the same sound with a 3KHz square wave added (to simulate engine tones) before the Doppler effect and level rise and fall were superimposed. They used the psychophysical Method of Pair Comparisons or Constant Stimulus Difference (CSD), with 1578 people and sound levels of 85 to 110 dB/SPL. They found a "fine but valid" distinction between

"loudness" and "disturbance" (The word cues used in their instructions to their subjects were "louder" and "more disturbing.>"). They also found that "the addition of discrete frequencies to a random noise (for equal overall SPL) has no effect on apparent loudness, but caused it to be judged some 2 dB more disturbing."

The idea that tones or bands of noise narrow enough to give a clear subjective perception of "pitch" could affect the "noisiness" of a sound, in particular a sound from an aircraft flyover, was studied by J.W. Little (1961, Ref. 2-27). In two experiments, he used the CSD method with 65 and 150 subjects, with sound levels of the order of 100 to 120 PNdB. The sounds used were jet engine noise and "pink" noise (noise having equal energy per octave) with and without discrete frequency ("spike") components. He found that the ability of spiked broad band noise to cause annoyance was related to the amplitude (relative to the background noise) and frequency of the spike, as well as the overall SPL of the noise, and that PNL "does not adequately assess the annoyance of spiked noise." He proposed a correction in PNdB, to be added to the PNL of the noise, proportioned to the relative amplitude of the spike and dependent on its frequency.

Where Copeland et al used jet engine noise with a superimposed square wave, Little used a more artificial sound and the results thus obtained may not be directly applicable to the much more complex sounds produced in a real aircraft flyover, in which there may be many tones of frequencies that may or may not be harmonically related, and that vary throughout the flyover in their perceptibility, which is mainly governed by the relative intensity of the tone and the "background" noise, which itself may have a complex and varying spectrum.

Kryter and Pearsons (1962, Ref. 2-20) used recordings of flyovers of aircraft of various sorts, with 23 subjects using the Method of Adjustment (MOA), with sound levels of 80-100 dBA, and judging "acceptability" or "disturbance." They concluded that PNL was somewhat more accurate than the other methods tried in evaluating noisiness or acceptability, but found a difference between jet aircraft noise

and piston aircraft noise, the PNL underestimating the relative noisiness of the jet noise by 3-4 dB. This they considered might be due to the presence of pure-tone components in the otherwise continuous spectra from some jet engines.

Because of the difficulty of controlling the presence and attributes of tonal components in recordings of real aircraft flyover noise, studies were continued using artificial noises. Such a study was reported by Kryter and Pearsons (1963, Ref. 2-21). Using the CSD method and sound levels of the order of 75 to 95 dB, they found that the presence of a tone in an octave band of noise increased annoyance by an amount that differed from that reported by Little (1961), possibly because Little used broad band noise as the "background" noise. They found a difference between "loudness" and "noisiness," and also that the duration of a sound influences the subjective response. In this report, the authors tabulated factors for calculating "noy" values for 1/3 octave bands of noise of given frequency and level. These tables were modified by Kryter and Pearsons in 1964 (Ref. 2-22) and 1965 (1965b, Ref. 2-24). PNL can be calculated using these tables and equation 2-2. In a manner analogous to the standardization of the A-, B- and C-weighting responses for measuring loudness (as opposed to its calculation), using the Fletcher and Munson equal-loudness curves, Kryter and Pearsons proposed an 'N'-weighting using the 40-noy equal-"perceived noisiness" contour for the measurement of PNL, to be expressed in dBN.

A further study by Kryter and Pearsons (1965, Ref. 2-23 and Ref. 2-24) again used single pure tones in octave bands of noise. In two tests, 21 and 20 subjects used the CSD method to compare sounds with levels of 40 to 115 dB. From the results, the authors proposed a correction procedure, in which a correction, depending on the frequency of the tone and its level with respect to the "background" noise, is added to the SPL of the band containing the tone. This corrected band level is then used in the computation of tone corrected PNL (PNLT) in the same way as PNL is calculated.

Pearsons and Horonjeff (1967, Ref. 2-46) reported a study using "noisiness" and "loudness" as descriptors, which consisted

of a laboratory test with 20 subjects judging recordings of aircraft flyovers and vehicle passbys of 70 to 100 dBA, using the Method of Numerical Category Scaling (NCS), and a field test with 42 subjects judging live flyovers and vehicle noises, at levels of 95 to 100 dBA, again using the NCS method. The sounds were evaluated objectively using PNL with the tone correction procedure of Kryter and Pearsons (1965), and a duration correction procedure of Pearsons (1966, Ref. 2-40). The authors reported that for the laboratory test, considering all stimuli, the correlation coefficient between subjective and objective evaluations was 0.78 for peak PNL and 0.79 for PNL corrected for tone and duration, a small improvement for the corrected measure. However, for the field study, considering only the aircraft noises, the correlation coefficient was 0.81 for both of these measures. Peak PNL was calculated by making traces of the sound level in each 1/3 octave band as it varied with time. The peak level for each band was measured and used in computing PNL. This calculates a measure often referred as "composite" peak PNL, and can be contrasted with maximum (or peak) PNL calculated by measuring the band level for each band at any one "instant" in time (often using 1/2-second sample time) and using these synchronous 1/3 octave band levels to compute PNL. This is repeated at each moment in time (every 1/2-second, for example) and the maximum value of PNL thus calculated is taken.

In 1967, Mabry and Little (Ref. 2-30) reported a laboratory study using recordings of aircraft flyovers presented at peak levels of 90 to 100 dB SPL (n.b. dB SPL by analogy with dBA, etc.) to 36 subjects who judged whether they would "complain" about the noise using a modified CSD method. The tone correction they used was one put forward by Little et al, reported by ISO/TC 43/WC 12 and 13, (1965, Ref. 2-13) in the fourth revised draft of the proposed Federal Aviation Administration noise certification criteria, in which a correction, based on frequency and relative level of the tone, is added to PNL. The duration correction used was proposed in the same document and is based on the duration in seconds of the PNL time history between the time at which the 1/2-second PNL first reaches

a value of 20 dB below its maximum value and the time at which it last decreases 20 dB below its maximum value (these are usually called the 20 dB-down points). Mabry and Little reported that the percentage of complaints increased with the tone correction but did not correlate with the duration corrections performed better than PNL. They also reported an interaction between the subjective effect of the tone and the SPL level of the noise, and suggested "that tone penalties should be assessed differentially as a function of SPL, "rather than solely on tone frequency and the relative levels of the tone to the "background" noise.

In 1968, Kryter (Ref. 2-15) summarized the then state-of-the-art and proposed a simplified tone correction procedure, based on that of Kryter and Pearsons (1965) and in which a correction, depending on the frequency and relative level of the tone, is added to the band level before calculation of PNL. He also proposed a method of correcting for the duration of a sound using an integration technique. All the energy in the sound during the time between the 10 dB-down points, measured using the PNL time history (concepts comparable to the 20 dB-down points discussed earlier), is summed to give an IPNL value (Integrated Perceived Noise Level).

Hecker and Kryter (1968, Ref. 2-10) reported a study using recordings of real and simulated flyover noises, with 20 subjects using the CSD technique to judge the acceptability of sounds with levels of 70 to 95 dBA. The objective measures used for comparison with the subjective responses included maximum values of dBA, dBC, dBN and PNL, and PNL with tone and duration corrections. The tone corrections used were those of Kryter and Pearsons [modified by Kryter (1968)] and a method developed by Little for the Federal Aviation Administration's noise certification program, proposed in a revised draft in February 1968. The duration corrections included an integration method proposed in the FAA document, based on Kryter (1968). Hecker and Kryter used a number of variants of this method, in which PNL, measured every 1/2-second between the 10 dB-down points and tone-corrected, in some variants, is integrated and compared with

a standard reference duration. Many variations on the duration correction have been used by different workers, the two main methods being based on the time duration between 10 dB-(20 dB- or other)down points (often referred to as the "estimated" duration correction) and on the integrated unit value (integrated over 10 dB, 20 dB, etc.) compared to a standard duration (of 10, 15, 20 seconds, etc.). This latter method is referred to as the "integrated" duration correction. The procedures can be applied to PNL, PLL, dBA, etc., with or without tone correction. Duration-corrected PNL is often designated PNLD; if the duration correction is the integrated version, this is sometimes referred to as IPNL. Tone-corrected PNL (PNLT) plus the integrated duration correction is EPNL (Effective Perceived Noise Level); PNLT plus the estimated duration correction is EEPNL.

Hecker and Kryter concluded that for a duration correction the "integration" method performed better than the "estimation" method. For the tone correction, they found PNL corrected by Little's method performed better than PNL alone or corrected by the method of Kryter and Pearsons.

Ollerhead (1968, Ref. 2-37) used recordings of aircraft flyovers presented at levels of 75 to 95 dB in a study in which 20 subjects judged the "noisiness" of the sounds, using the CSD method. The objective measures used included overall SPL, A-, B-, C- and N-weighted SPL, Stevens' and Zwicker's phons (PLL) and PNL, with a tone correction from a proposed draft of the FAA procedure (December 1967), together with other, non-standard weighting functions. He found that the tone correction improved the correlation coefficient for maximum PNL from 0.880 to 0.900; when the duration correction was applied, the correlation coefficients were 0.837 for PNLD and 0.846 for tone-corrected PNLD (EPNL).

Hinterkeuser and Sternfeld (1968, Ref. 2-12) used synthesized flyover-type sounds, using actual aircraft sounds as a starting point and simulating predicted V/STOL noise. Eighty-two subjects judged sounds at levels of the order of 110 PNdB using the CSD technique. Peak values of PNL were calculated and corrected with tone and duration corrections (FAA, revised draft, August 1966) using 1/1

octave-band data. The authors reported that, when considering all the signals, the average calculated level difference in dB between comparison and standard sounds at subjective equality was -1.1 for PNL, -0.37 for PNLT, +1.0 for PNLD and +1.45 for EPNL. The standard deviation of the results was 4.2 for PNL, 2.73 for PNLT, 2.82 for PNLD and 2.9 for EPNL. They stated that their statistical evaluation "does not show any greatly significant effects of the corrections." However, for the group of sounds in which pure tone components were most strongly evident, "a significant improvement in correlation is indicated by inclusion of the pure tone correction factor."

Pearsons et al (1968, Ref. 2-47), and 1969, Ref. 2-45) made a particular study of the effects on human response of tones in noise. They used PNL with tone corrections of Little, and of Kryter and Pearsons, both defined in Pearsons et al (1969), and studied a variety of tonal stimuli, including single tones, multiple tones and modulated tones, superimposed on broad band or octave bands of noise at levels of 70 to 100 dB. Though the broad band noises used as "background" included noise weighted in frequency to simulate the spectrum of noise from a turbojet, the stimuli used in this study were very much simplified compared to the complexities of flyover sounds. With these simplified signals, the authors demonstrated "a clear difference between the results obtained with noisiness and loudness" as the descriptor in the subjects' instructions but noisiness "without further definition may have been interpreted as 'simply loudness'." They reported that "the judgment tests clearly confirm the need for a discrete frequency correction" for these stimuli, but found no significant difference between the two tone correction procedures. They found differences in the results between broad band noise and octave bands of noise as background.

In 1968, Sperry (Ref. 2-56) reported a standard method of calculating EPNL and defined the tone and duration corrections to be used. These methods were incorporated in the FAA Federal Aviation Regulation Part 36 (FAR-36) for the standardization of noise measurement for aircraft certification legislated in July 1968. The tone detection and correction methods used in FAR-36 followed Sperry until

March 1978 when an amendment was made in the tone correction calculation, and the possibility of "band-sharing" of tone between adjacent 1/3 octave bands was considered.

Pearsons (1968, Ref. 2-41) reported more studies of tones superimposed on broadband noise, in which the time patterns were varied. He stated that "varying the duration of the tone provided little change in the judgment results compared with those results where the duration of the tone was comparable with the duration of the noise. Also, the time at which the tone peak occurred did not seem to affect greatly the judgment results." He concluded that the PNL with a tone correction predicted the noisiness of the stimuli employed, though for complex stimuli varying in both tone content and duration, a duration correction was also necessary.

Kryter et al (1968, Ref. 2-18, and 1970, Ref. 2-19) reported a large study using live aircraft flyovers of levels of 80 to 120 PNdB, which were judged by 96 subjects using the CSD technique for "acceptability." The authors concluded that integrated units, such as EPNL, better predicted judged perceived noisiness than did maximum or peak units, and that "tone-corrections did not contribute significantly for these noises to the predictive accuracy of the various physical units used." Among the procedures studied were a change in the calculation of PNL, to provide weightings more proportional to the critical bandwidths of the ear, and a proposed weighting contour in line with these changes, designated dBD. The tone corrections studied were those of Kryter and Pearsons (1965) and the FAA (Sperry, 1968). The average differences in dB between reference and comparison aircraft noises when judged equally unacceptable were reported as -0.8 for EPNL (with Kryter and Pearsons tone correction), -1.2 for EPNL (with FAA tone correction), -1.6 for PNLD, 0.1 for maximum PNLT (Kryter and Pearsons), 0.3 for maximum PNLT (FAA), and -0.7 for maximum PNL, showing only slight differences. The standard deviations for these units were 3.7, 3.8, 3.6, 5.8, 4.8 and 5.0 respectively, showing some improvement with the use of the duration correction but little effect from the tone correction

on the duration-corrected unit.

Little and Mabry (1968, Ref. 2-28, and 1969, Ref. 2-29) used 35 calculation procedures in a study in which recordings of aircraft flyovers were presented at levels of 90 to 100 EPNdB to 36 subjects who rated them for annoyance or objectionability using the CSD method. The authors looked at a number of tone-correction methods including those of the FAA (5th revised draft, 1967; see Sperry, 1968), Kryter and Pearsons (1965), and the FAA (4th draft 1967; see ISO/TC-43, 1967), as well as two methods of detecting the presence of a tone from a 1/3 octave spectrum of the noise. This latter problem is clearly an important consideration in the evaluation of a tone-correction procedure; if the quantity of the tone correction is a function of the frequency and relative level of the tone compared to the background noise, then it is necessary that the tone detection method (a mathematical manipulation of the 1/3 octave band levels) should not identify non-existent (imperceptible) tones, should assign identified tones to the correct 1/3 octave band, and should estimate the levels of the tone and the background noise accurately. A number of detection procedures exist, including that used by Kryter and Pearsons (1965) and Kryter (1968), the "four-band average" or "two-pass averaging" method (used in the 4th draft of the FAA proposals, 1967) and the "slope" method (used by Sperry, 1968). Little and Mabry studied these last two methods. They found, for the most part, that the "four-band" method of detecting tones resulted in greater precision than the "slope" method, that the 5th draft FAA tone correction was the superior procedure in computing PNLT, and that the duration correction degraded its performance.

Pearsons (1969, Ref. 2-42) reported a study using single tones combined with broad band noise spectra and turbofan engine runup noise shaped to simulate a flyover, with varying time histories (some tone peaks not coincident with broad band noise peaks). Twenty subjects used the CSD method to judge the more objectionable

or disturbing noises; levels used ranged from 70 to 95 dB. The author found that tone and duration corrections appeared to be additive for these stimuli (i.e., no interactions were found), and that the tone and duration corrections used [Kryter and Pearsons, (1965) and Pearsons (1966)] were adequate in predicting the noisiness of these laboratory-generated noises. However, the single real life engine noise used "appeared to be judged consistently noisier than the corrected PNL would predict."

Pearsons and Bennett (1969, Ref. 2-43, and 1971, Ref. 2-44) used a modified version of the CSD technique, called PEST (Parameter Estimation by Sequential Testing) in which the presentation levels are dependent on past responses from the subject and are under computer control. The signals used were shaped random noise, varying in spectrum and time-history, superimposed on some of which were pure tone components. Tone corrections considered were those of Kryter and Pearsons (1965), and the FAA (Sperry, 1968). The authors reported that tone corrected perceived noise level was an improvement over the uncorrected perceived noise level, e.g., 1 dB for $PNLT_{CKP}$ (Kryter and Pearsons tone corrected, "composite" peak PNL) and 2 dB for $PNLT_F$ (FAA tone corrected, maximum PNL). For signals with tonal components and large variations in duration "both tone and duration corrections were required to provide the greatest improvement in noisiness prediction measures." However, in a test using recordings of various types of aircraft flyovers at levels of 70 to 80 dBA, the stimuli "did not indicate large differences between various measures." The authors reported that standard deviations were of the order of 1.5-2.5 dB for all units except overall SPL which was noticeably worse.

Adcock and Ollerhead (1970, Ref. 2-1) tested 32 subjects with recorded flyover sounds, using the CSD method with sound levels of 80 to 110 dBA and judging "noisiness" and "disturbance." The tone correction procedure used was that of Sperry (1968). The authors found that the efficacy of the duration correction depended on the type of sound, i.e., it improved the correlation of subjective

and objective measures for sounds from STOL aircraft but degraded it for CTOL aircraft sounds. The tone correction included in the PNLT and EPNL scales "appears to degrade the performance of the PNL scale for the particular class of sounds studied." This they attributed to the presence of only three (out of 60) signals having evident high frequency tone components. The tone correction procedure was detecting and correcting some of the signals for low frequency tones that were not subjectively present; these "tones" were reported to be "essentially harmonic components of pulsatile propeller and exhaust sounds, and as such are not really heard by the observer as pure tones, in the same sense as are high frequency compressor tones." The authors proposed that "a low frequency cut-off might be imposed in the EPNL procedure so that tones detected below a certain frequency, probably in the neighborhood of 500 Hz, are ignored."

In June 1970, the International Organization for Standardization published a recommendation for calculating EPNL (ISO R507, Ref. 2-2) which agrees with FAR-36 (prior to March 1978) in its method methods of calculating PNL [using tables or the mathematical method of computing PNL using equations developed by Pinker (1968, Ref. 2-48)], in computing the duration correction (based on integration over the 10 dB-down range), and in computing the tone correction.

Kryter (1970, Ref. 2-16) reviewed previous literature, and reanalyzed results from previous experiments. He proposed a change in the computation of PNL, by weighting sound energy below 355 Hz better to account for the critical bandwidth of the ear. He used the Kryter and Pearsons (1965) tone correction and that of Sperry (1968). He found that durational information (between the 10 PNdB-down points) significantly improved the predictive accuracy of PNL and Effective units (e.g., EPNL) were appreciably better than Estimated Effective units. He reported that the maximum PNL scale was sometimes degraded by using a tone correction, which he attributed to the signal containing an audible tone during part of its time history (and thus causing annoyance) but the tone not being

present at the moment of maximum PNL (and thus not being adequately taken into account). This effect would be less apparent for the integrated units; hence EPNL is slightly more accurate when tone-corrected than not tone-corrected.

Wells (1970, Ref. 2-61) used the MOA technique with 35 subjects rating aircraft engine sounds reproduced as recorded or after filtering. The sounds were presented at levels of the order of 90 PN_dBT, using the tone correction procedure of FAR-36 (1970) (see Sperry, 1968). The author reported that PNLT seemed to overrate the subjective importance of isolated tones by two to four decibels, but underrated the subjective importance of multiple pure tones by two to four decibels. He also found that a method of measuring annoyance level (ANL) that he had previously proposed (in 1969, Ref. 2-59 and 2-60) rated actual engine spectra better than did PNLT, or any of the other measures he had considered.

Langdon et al (1970, Ref. 2-26) presented 41 subjects with recordings of aircraft flyovers, which they judged for "acceptability," using the CSD method. The tone and duration corrections used were those given by Sperry (1968). The authors reported that maximum PNLT, maximum PNL, EPNL, PLL and dBD provided the best agreement between scale data and judgment data of the scales evaluated, but did not differ from each other to a statistically significant extent.

In 1971, the Boeing Company (Ref. 2-6) published a report of a study in which 180 subjects judged live aircraft flyovers and USASI noise signals using the psychophysical methods of CSD and Magnitude Estimation (ME). Signal levels were of the order of 80 to 115 PN_dB and subjects rated them for "annoyance," "disturbance," and "noisiness." It was reported that the use of a tone correction did not "improve the basic procedures" (both PNL and PLL were studied); also the duration correction "decreased the relationship" between subjective and objective measures. The corrections used were those of FAR-36 (1969) (see Sperry, 1968) and of the third draft of the FAA proposed regulations (the tone being identified by both the

"four-band averaging" and the "slope" technique).

Ollerhead (1971, Ref. 2-38, and 1973, Ref. 2-39) presented flyover recordings at levels of 85 to 115 dB to 32 subjects, who judged them for "noisiness," "objectionableness," "disturbance," and "unwantedness" using the CSD method. The tone correction procedure used was that of the FAA (FAR 36, 1969; see Sperry, 1968). The author reported that the integrated duration correction had a "beneficial effect on the performance of the scales," but the tone correction did not prove "a particularly beneficial measure" since, in general, its application caused both PNL and PNLD "to become less consistent evaluators of perceived level." He observed that large tone corrections were applied by the method to signals with low frequency components in a manner not consistent with perceived effects, and recommended that apparent "tones" below 500 Hz should be ignored as an interim measure until the manner of detection of tones could be improved.

The Society of Automotive Engineers Research Project Committee R-6 (SAE R-6) reported on a study (1971, Ref. 2-52) in which 30 people judged seven flyover recordings in each of two studies performed at different laboratories. In one case, levels of the order of 90 PNdbT were used; in the other, levels of 80 PNdbT; in both "acceptability" and "annoyance" were judged. The tone correction studied was that in FAR-36 (see Sperry, 1968). The results showed that integrated ANL, EPNL, maximum dBA and maximum dBD "can be expected to exhibit reasonably small standard deviations when compared with juror ratings," and that the multiple pure tone correction did not improve the correlation of ANL with juror ratings, as had previously been found.

Kryter (1972, Ref. 2-17) presented a reanalysis of earlier experiments using aircraft and simulated aircraft noises, and concluded that the FAA tone correction (Sperry, 1968) did not differ significantly from that of Kryter and Pearsons (1965). The average standard errors in dB reported for PNLD, EPNL (tone corrected

Kryter and Pearsons) and EPNL (tone corrected FAA) were 3.82, 2.84 and 3.12, respectively, and for the equivalent maximum values [PNL, PNLT (Kryter and Pearsons) and PNLT (FAA)] were 3.99, 3.95 and 3.79. The author reported that EPNL is significantly better than maximum PNL.

Goulet and Northwood (1973, Ref. 2-9) reported on a study that used broad band noise with an added pure tone and indicated "the A-weighted level remains an adequate (slightly conservative) rating number even when pure tones are present."

Powell (1973, Ref. 2-49) used synthesized turbofan STOL aircraft flyover sounds with recordings of CTOL aircraft in a study in which 20 subjects judged "annoyance" using the ME method, with signal levels of 65 to 95 PNdB. The tone correction was that of FAR-36 (1969)(see Sperry, 1968). The author reported that "the use of tone corrections did not improve the accuracy of the scaling units considering all of the aircraft sounds" used in the study.

The SAE R-6 committee published a report (1973, Ref. 2-53) on two parallel experiments in which 60 subjects judged real and synthesized aircraft flyover sounds presented at levels of 60 to 90 dB in one case and 80 to 110 dB in the other, using the NCS technique and rating them for their "unpleasantness." In one study, tones were detected by the "four-band average" technique and in the other the "slope" method was used. Corrections were computed by FAR-36 (see Sperry, 1968). Linear regression correlation coefficients for maximum PNL, maximum PNLT, and EPNL were 0.895, 0.935 and 0.881 in one study and 0.949, 0.903 and 0.901 in the other. Differences are therefore neither large nor consistent. In October 1973, the Society of Automotive Engineers issued a standardized method of calculating EPNL as ARP 1071 (approved by the American National Standards Institute in July 1973 as ANSI S6.4-1973) (Ref. 2-3). The method of calculating PNL is defined in SAE ARP 865A, and is substantially that used in FAR-36. The duration correction is based on integration of the signal over the 10 dB-down range, as is FAR-36.

The tone correction procedure however differs from that of FAR-36; the detection method is different and the correction values also differ somewhat, in that tone corrections of less than one dB are equated to zero. With a slight change in relative 1/3-octave band levels such as can occur with reanalysis of a flyover signal, the tone correction can thus change from 1.0 dB to 0.0 dB, which may be a significant effect. With FAR-36 (up to March 1978), tone corrections of less than one dB are ignored for 1/3-octave band center frequencies of less than 500 Hz or more than 5 KHz. Between these two limits, tone corrections are allowed as low as 0.5 dB, below which they are ignored.

Mabry and Perry (1973, Ref. 2-31) evaluated four psychophysical methods (ME, NCS, MOA and CSD) using recordings of flyovers and artificial noises at levels of 80 to 100 PNdB, played to groups of 16 or 24 subjects who rated them for "annoyance." The authors reported some interesting conclusions as to the comparative methods and general techniques in psychoacoustical testing. The effects of tone and duration corrections on PNL were reported as small, though the tone correction procedure tended to reduce slightly the correlation coefficients.

The SAE R-6 committee published a report in 1973 (Ref. 2-54) of a study in which 24 subjects were required to rate 48 flyover recordings, played at levels of 70 to 80 dBA, with the NCS technique, using "annoyance" as the cue. Only two units were compared, maximum dBA and EPNL (using FAR-36 tone and duration corrections). Though the study described in this report was not as complete as originally envisaged, "the results obtained clearly (showed) that neither EPNL nor dBA (could) effectively rate the group of aircraft flyover sounds used." It was noted that "dBA (was) consistently more effective than EPNL."

Kryter et al (1974, Ref. 2-25) used recordings of synthesized and real aircraft flyover noises which were played to 72 subjects

at levels of 65 to 90 dBA; judgments were made with both the CSD and ME techniques of the "noisiness" or "unacceptability" of the sounds. The tone corrections used were those of Sperry (1968) and Kryter and Pearsons (1965); they were found to be of some utility with regard to improving the predictions of the subjective judgments, but their effects were rather small. Standard deviations for maximum PNL, the two tone-corrected PNL values and their duration-corrected equivalents were given as 1.99, 2.03, 2.48, 2.64, 2.36 and 2.29, which would rate maximum PNL as the best measure, though differences are minimal.

Powell and Rice (1975, Ref. 2-51) reported on an investigation of subjective response to aircraft noise in a traffic noise background. Twelve subjects judged flyover recordings at levels of 50 to 65 dBA for "annoyance" using the NCS technique. The authors looked at various weightings, PLL and PNL with tone and duration corrections, and found "there were no major or consistent differences between the predictive abilities of the various rating scale units."

Berglund et al (1975, Ref. 2-5) used a laboratory study to test if subjects would rate aircraft noise for "loudness," "noisiness" and "annoyance" consistently and differently. These descriptive terms were carefully defined to the subjects by the experimenters: "loudness" as the perceptual aspect of the noise that is changed by turning the volume knob on a radio set," "noisiness" as "the quality of the noise" (for example, "the sound from a jackhammer may be more or less noisy than that from a motorbike even if they are considered equally loud" and similarly, "music may be loud but still not perceived as noisy"), and "annoyance" as "the nuisance aspect of the noise experienced in an imaginary situation phrased as: 'After a hard day's work, you have just been comfortably seated in your chair and intend to read your newspaper.'" The authors state that Kryter's concept of perceived noisiness "is ambiguous and covers both the noisiness and annoyance concepts introduced here in an attempt to differentiate between them." This statement well illustrates the difficulty of this aspect of psychoacoustical testing: different experimenters have different concepts of what a word means; they may try to impose

their interpretations on their subjects whose reactions may be thereby influenced or confused. Observers who have equated "annoyance" with "noisiness" have ignored the contextual nature of "annoyance" which Berglund et al tried to define by using the analogy of resting with a newspaper after a hard day's work. Such imaginative requirements may be difficult for a subject to respond to in a laboratory test; though their results may be consistent there appears to be no evidence as to how well the laboratory results truly reflect the imagined situation. It is a matter of continuing controversy as to what subjective responses are being measured in a laboratory test; perhaps it would be justifiable to leave the subjects to draw their own conclusions as to the word cues used ("annoyance," "objectionable," "unacceptable," etc.) if these are words that are in common use and that any individual is likely to comprehend and use consistently (even if in his own interpretation). The ideal test might be considered to be one taking place outside the laboratory in an unforced natural setting.

MAN, Inc. (1975, Ref. 2-33) presented 35 subjects with real and synthesized aircraft flyover noises, including helicopter sounds, at levels of 55 to 70 dBA; judgments of "annoyance" were made using the ME technique. The units studied were PNL and dBA, with tone and duration corrections from FAR-36 (1969) (see Sperry, 1968) and Stevens' Mark VI and Mark VII PLL. The results of an analysis of variance showed that all ten objective units had highly significant F-ratios, indicating that no unit predicted subjective response well. The lowest F-ratio was 24.58 for PNLD; that for EPNL was 26.10, for maximum PNLT was 47.61, and for maximum PNL was 46.42. Thus the tone correction has little effect but the duration correction reduced the F-ratio somewhat.

MAN, Inc. (1976, Ref. 2-34) reported another study using aircraft flyover signals with helicopter and simulated VSTOL recordings at levels of 55 to 80 dBA presented to 24 subjects who rated them for "annoyance" using the ME technique. F-ratios for PNL, PNLT, PNLD

and EPNL were reported as 18.07, 20.04, 5.52 and 6.63, considering all the signals used in the study. Thus the duration correction caused an improvement in the unit's ability to predict subjective reaction, whereas the tone correction degraded the unit's performance.

Powell (1977, Ref. 2-50) presented flyover recordings in two tests to 96 and 32 subjects at levels of 65 to 95 dBA (representing levels that would be heard outdoors) and 40 to 85 dBA (levels that would be heard indoors); judgments of "noisiness" (qualified by the words "unwanted," "objectionable," "disturbing," "unpleasant") were made using the NCS technique. Of the units investigated, the most consistent in predicting the noisiness for all aircraft were maximum dBA, Stevens' Mark VII PLL (with and without duration corrections) and EPNL [using the FAR-36 procedures; see Sperry (1968)]. Maximum PNLT was found to be the least consistent scale. Correlation coefficients for the "outdoor levels" experiment were 0.962 for maximum PNL, 0.974 for EPNL and 0.942 for maximum PNLT; for the "indoor levels" experiment using estimated outdoor levels, the correlation coefficients were 0.958, 0.971 and 0.933, respectively. The author reported that the tone correction procedure added a 1.2 dB tone correction to one of the stimuli (a recording of the Concorde S.S.T. take-off) in which no tonal component was audible. "Closer examination of the 1/3 octave band and 1/2 second time histories revealed that tone corrections ranging from 0.0 to 2.4 dB occurred (in this recording) randomly in both time and frequency of the 1/3 octave bands between 500 Hz and 1000 Hz." The tone corrections in noises with true tonal qualities, e.g., the DC-8 turbofan landing noise, were not nearly so random. It would therefore appear that the tone correction procedure was incorrectly detecting nonexistent tones, resulting in unnecessary tone corrections.

Mabry and Sullivan (1978, Ref. 2-32) presented real and synthesized flyover recordings (at levels of 55 to 80 dBA) to 60 subjects who rated them for "annoyance" using the ME technique. The units used included PNL with tone and duration corrections from FAR-36 (see Sperry, 1968). The standard deviations of the subjective dB

results* for maximum PNL, maximum PNLT, PNLD and EPNL were respectively 2.0, 1.9, 2.2, and 2.1 dB. The tone correction produced a very small improvement, while the duration correction slightly degraded the measure. The differences were not significant however.

In March 1978, the Federal Aviation Administration produced a revised version of its procedure for the calculation of EPNL (Ref. 2-4). The tone detection method was unchanged, but the correction values were altered to give correction factors down to 0.0 dB for very low relative intensity tones, which protrude above the background by less than 3 dB. A procedure to compensate for "band-sharing" was also introduced; if a tone falls between two 1/3 octave bands, its energy will be shared between those two bands which may result in the detection of an artificially low-intensity tone. However, in an aircraft flyover recording, the Doppler effect causes the frequency of a tone to change, so a tone that is shared between two bands during one time-sample will fall into only one band during nearby time samples. It is therefore stated in FAR-36 §836.5m that after the maximum value of PNLT is identified, the frequency of the largest tone correction factor must be identified for the two preceding and the two succeeding 500-millisecond time intervals. If the largest tone correction for maximum PNLT is less than the average value of the maximum tone corrections for those five consecutive time intervals, that average value of the maximum tone correction must be used to compute a new value for maximum PNLT.

For some flyovers with indistinct tonal components, the frequency band containing the maximum tone correction may vary in successive 1/2 second intervals (as with Powell's recordings of the Concorde in Powell, 1977) in a manner that is not due to Doppler effect; with Doppler effect, the frequency of the tone and therefore the band number of the tone correction will decrease steadily for a normal flyover. This provision of FAR-36 has thus been interpreted to mean that, for 1/2 second samples adjacent to the peak, the highest tone

*For an explanation of subjective dB, see Section 4 of this report.

correction in the bands within one octave of that band giving the maximum tone correction at maximum PNLT shall be used to calculate the five-sample average tone correction for comparison with the tone correction at maximum PNLT.

McCurdy and Powell (1979, Ref. 2-36) reported on a study that used synthesized flyover sounds at levels of 65 to 95 dBA which were presented to 48 subjects who judged them for "annoyance" using the NCS technique. When using results from all stimuli together, EPNL performed the best [using the duration and tone correction procedure of FAR-36, see Sperry (1968)]. When results from signals with high tonal content were separated from those without tones, EPNL is less effective. Thus it appeared that the tone correction used in EPNL "aids in comparing the annoyance of stimuli with distinctly different tonal content but may slightly degrade the prediction ability when used in comparing stimuli of similar tonal content." The authors of this publication also reported that an analysis of variance of their results showed that annoyance is significantly affected by the tonal content of a noise. They indicated two areas of possible improvement in the PNLT tone correction method. "First, a change in the procedure to account for the apparent interaction of tonal content and sound pressure level," (which their analysis had shown) and "second, a modification to the procedure to prevent the application of a tone correction to stimuli which contain no tones." They stated that "the prediction of the effects of tonal content appears to be the largest remaining source of variation in the prediction of overall annoyance response."

Scharf and Hellman (1979, Ref. 2-55) published a report in which they reanalyzed data from many previous studies. Different studies have investigated different units with different statistical techniques; this was an attempt to coordinate results to discover if any consistent findings could be formulated. The tone corrections used were those of FAR-36 (1969) (see Sperry, 1968), of Kryter and Pearsons (1965), and one tentatively proposed by S.S. Stevens (1970, Ref. 2-58).

The authors concluded that "a detailed analysis of over 500 spectra with and without tonal components provided little evidence of the need for a tone correction." However, some of the studies used by the authors required judgments of loudness or evaluative judgments at levels below 80 dB. These factors may have reduced the likelihood of the data showing any effects of tonal components. The small effects of tonal components in the group of studies analyzed by Scharf and Hellman precluded any definite conclusions about the relative merits of the tone correction methods considered. However, none of the three methods improved the effectiveness of the units to which they were applied; "the variability and the discrepancy between calculated and judged level either remained the same or increased." The authors concluded that "data are needed on a large enough set of sounds with and without tonal components to permit adequate evaluation of tone correction procedures."

May and Watson (1980, Ref. 2-35) published another report in which data from previous studies were reanalyzed. The tone corrections considered were those of FAR-36 (March, 1978), ARP 1071 (1973), and Kryter and Pearsons (1965). The authors point out the difficulties of detecting tones from 1/3 octave band spectra and recommend consideration of narrower band analyses. They concluded that EPNL calculated using any of the tone correction procedures was equally effective; the standard deviation for EPNL (FAR tone correction) was 1.95, for EPNL (ARP tone correction) was 2.12, for EPNL (K&P tone correction) was 2.48, for PNLD (no tone correction) was 1.70. Thus PNLD with no tone correction performs slightly better than any of the other units. The authors also investigated data from a reduced number of flyover signals than were used to calculate the above results. They found a correlation between the size of the tone correction and the shape of the spectrum for flyovers in the total data base; to reduce any skewing of data due to the interaction of these factors, they selected a reduced data base in which this correlation was minimized. For the reduced data

base, the standard deviations were calculated to be 1.95 for EPNL (FAR tone correction), 2.02 for EPNL (ARP tone correction), 1.57 for EPNL (K&P tone correction) and 1.89 for PNLD. Again differences are small, though there is a tendency for the Kryter and Pearsons tone correction method to improve the relationship between subjective and objective measures.

2.3 DIRECTION FOR THIS STUDY

This survey of previous work in the field of tonal corrections to be applied to measures of aircraft noise has shown a lack of clear evidence in any direction. Most studies have shown little or no effect either positive or negative from the application of such corrections. The studies where tonal corrections had had most efficacy have been ones in which artificial (when compared to aircraft noise) sounds have been used. The sounds used have been either too homogeneous or too heterogeneous to show any clear effect.

The most consistent conclusion that has been drawn by previous workers is the need for more research; one comment that has been made is the inaccuracy of the tone detection procedures, especially in consideration of low frequency components. The findings of Powell (1977) of random tone corrections of from 0.0 to 2.4 dB indicates a definite lack of consistency in the procedures in use.

The use of the PNL method of describing "noisiness" of aircraft noise is well established, although it has been questioned by some researchers, including Wells and Ollerhead. It thus seemed reasonable to investigate presently existing tone correction procedures that have been standardized for use with PNL. The procedures chosen for the study described in Sections 3 through 6 of this report were those standardized by the FAA (FAR-36, March 1978), the SAE (ARP 1071, October 1973), and the ISO (ISO R-507, June 1970).

The three standard procedures for detection and correction for tones bear strong resemblances to each other. FAR-36 uses the same detection method as ISO R-507, whereas ARP 1071 uses a different

procedure. The correction factors to be applied to the detected tone are functions of frequency and relative level, and are the same for all three methods, except for low level tones. FAR-36 corrects for detected tones, however small their relative levels, ISO R-507 corrects for tones that protrude by 3 dB (the difference in dB between the 1/3 octave band level of tone plus noise and the predicted level of the 1/3 octave band of noise alone) or greater, and ARP 1071 has the most abrupt transition (between 0.0 dB and 1.0 dB tone correction), with ISO R-507 lying between.

These three methods divide the frequency scale into three sections: mid-frequencies (500 Hz to 5 KHz), in which the tone correction is highest, and low- and high-frequencies (<500 Hz or >5 KHz) which are given tone corrections half that of the mid-frequency band.

Because of the strong similarity between the three standardized methods of tone correction, it was decided for comparison to investigate two very different procedures, that of Little et al reported in the Fourth Draft of the FAA Procedures (ISO/TC-43, 1967) and that of Kryter and Pearsons (1965), extended to the 8 KHz region by May and Watson (1980). The Little correction was used with the detection procedures of the standardized methods. The Kryter and Pearsons correction differs from all the others considered in that it is applied to the band SPL before computation of PNL. Kryter and Pearsons proposed their own detection procedure which was followed in this study; use could, however, be made of their correction factors with only one of the other detection methods. Both the Little and the Kryter and Pearsons correction values are more complex than any of the standardized methods; they vary with frequency continuously instead of having the frequency dimension divided into three discrete sections. Relative correction factors of the five procedures are shown in Figures 2-1, 2-2, and 2-3.

PNL is the standardized frequency-dependent level correction method for use with aircraft noise; however, a number of studies

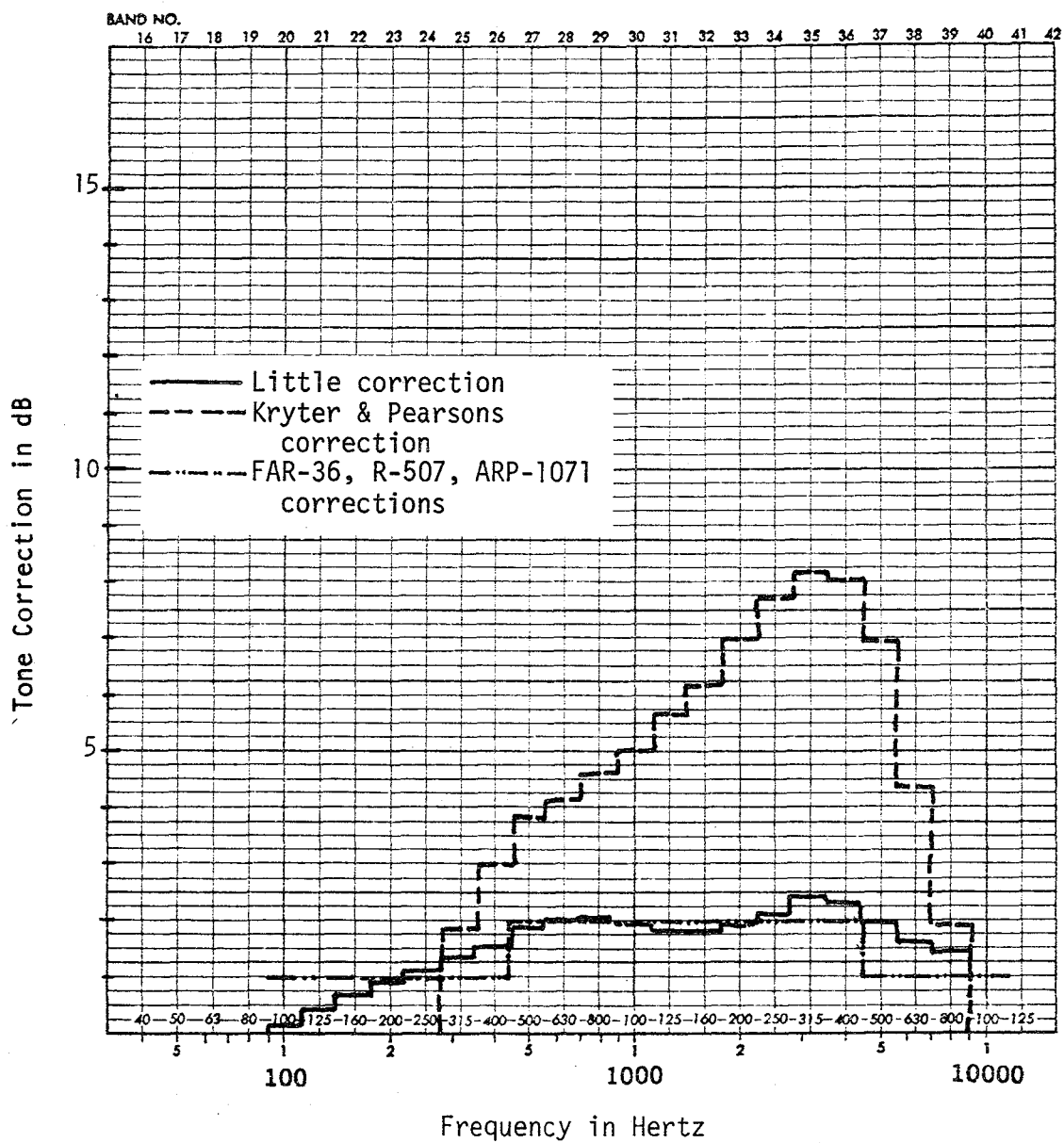


Figure 2-1. Difference Between 1/3-Octave Band Levels of (Tone plus Noise) and Noise Alone is 6.0 dB.

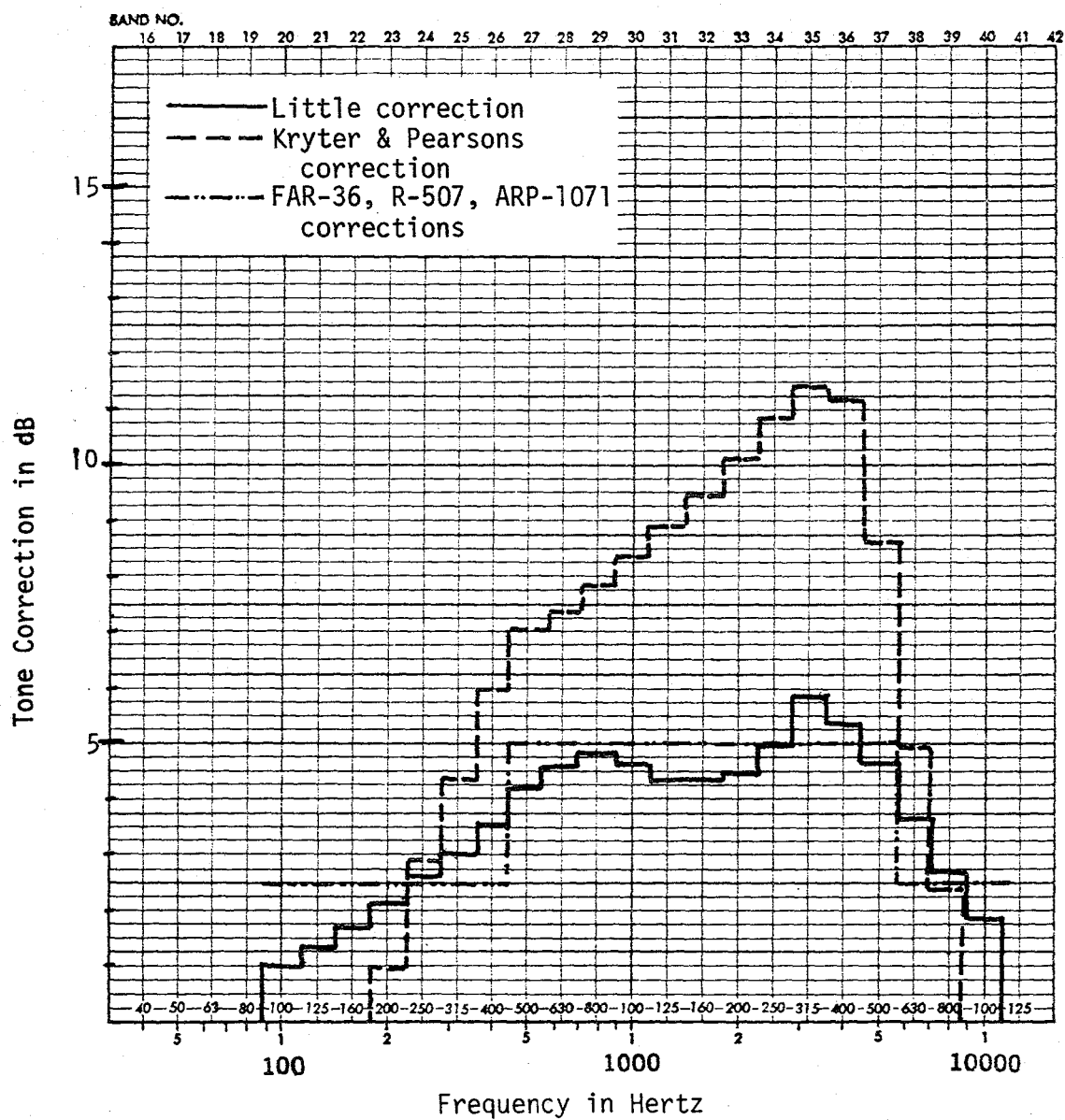


Figure 2-2. Difference Between 1/3-Octave Band Levels of (Tone plus Noise) and Noise Alone is 15.0 dB.

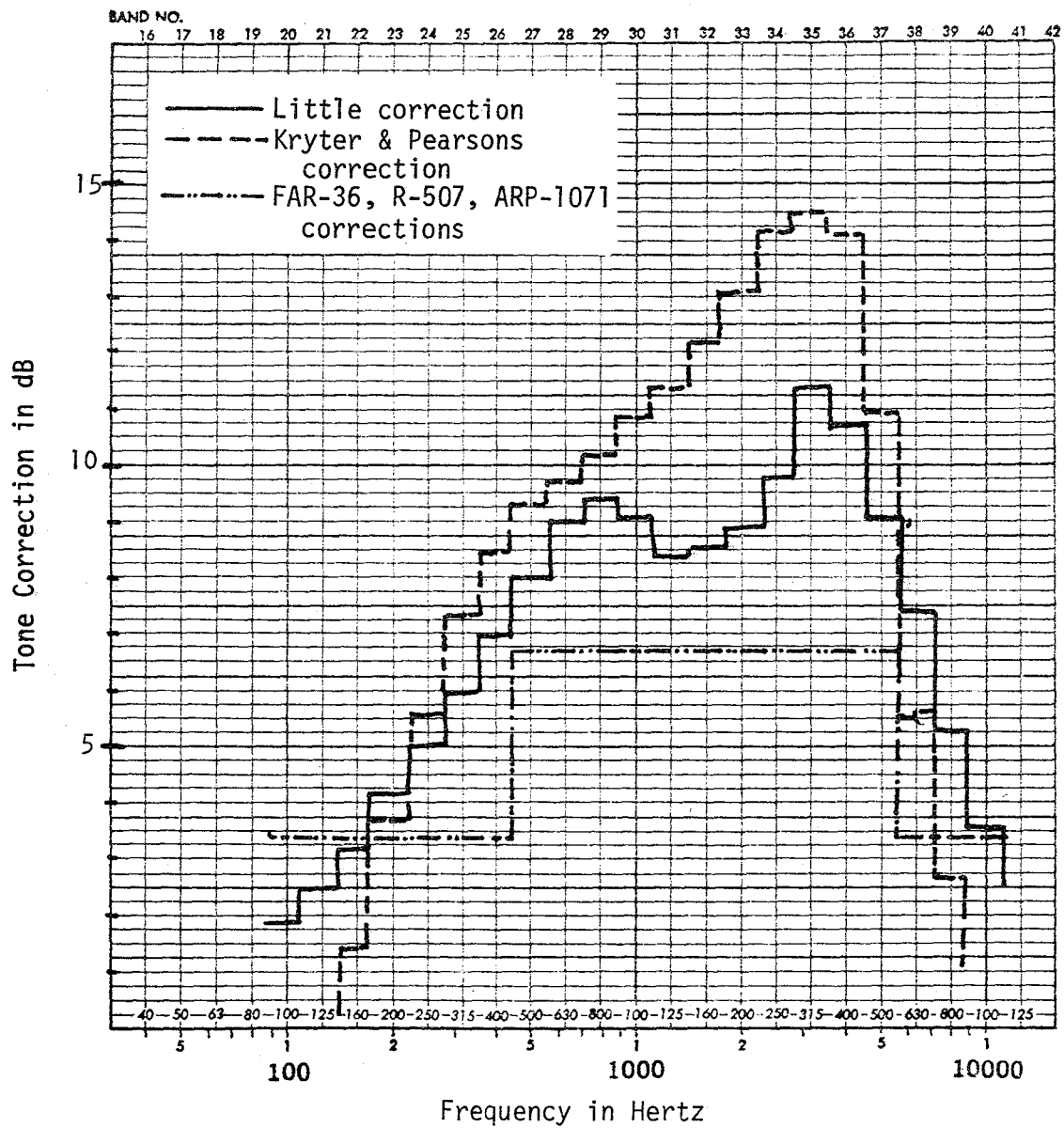


Figure 2-3. Difference Between 1/3-Octave Band Levels of (Tone plus Noise) and Noise Alone is 25.0 dB.

have reported that dBA is as adequate as PNL in predicting human response. Therefore it was decided to include both weighting functions in this analysis.

Two studies were performed, one at levels that could be heard outdoors near an airport and one using levels to be heard indoors in a similar location. This was intended to contribute data to the possible dependency of the tone correction on absolute signal level. The different levels were achieved by filtering the signals through a standard construction house wall; this not only shifted the signal levels but also altered their spectra. A comparison of the USASI noise used in the experiment as a standard signal for the two facilities is shown in Figure 2-4, to illustrate this spectrum change.

The spectrum change altered the signals in such a way that, where for the "outdoor levels" test the predominant tone frequencies were of the order of 2-3 KHz, for the "indoor levels" test, the dominant spectral peak was of the order of 160 to 200 Hz. This enabled some investigation of the possible overcorrection or misidentification of low frequency tonal components noted by other workers.

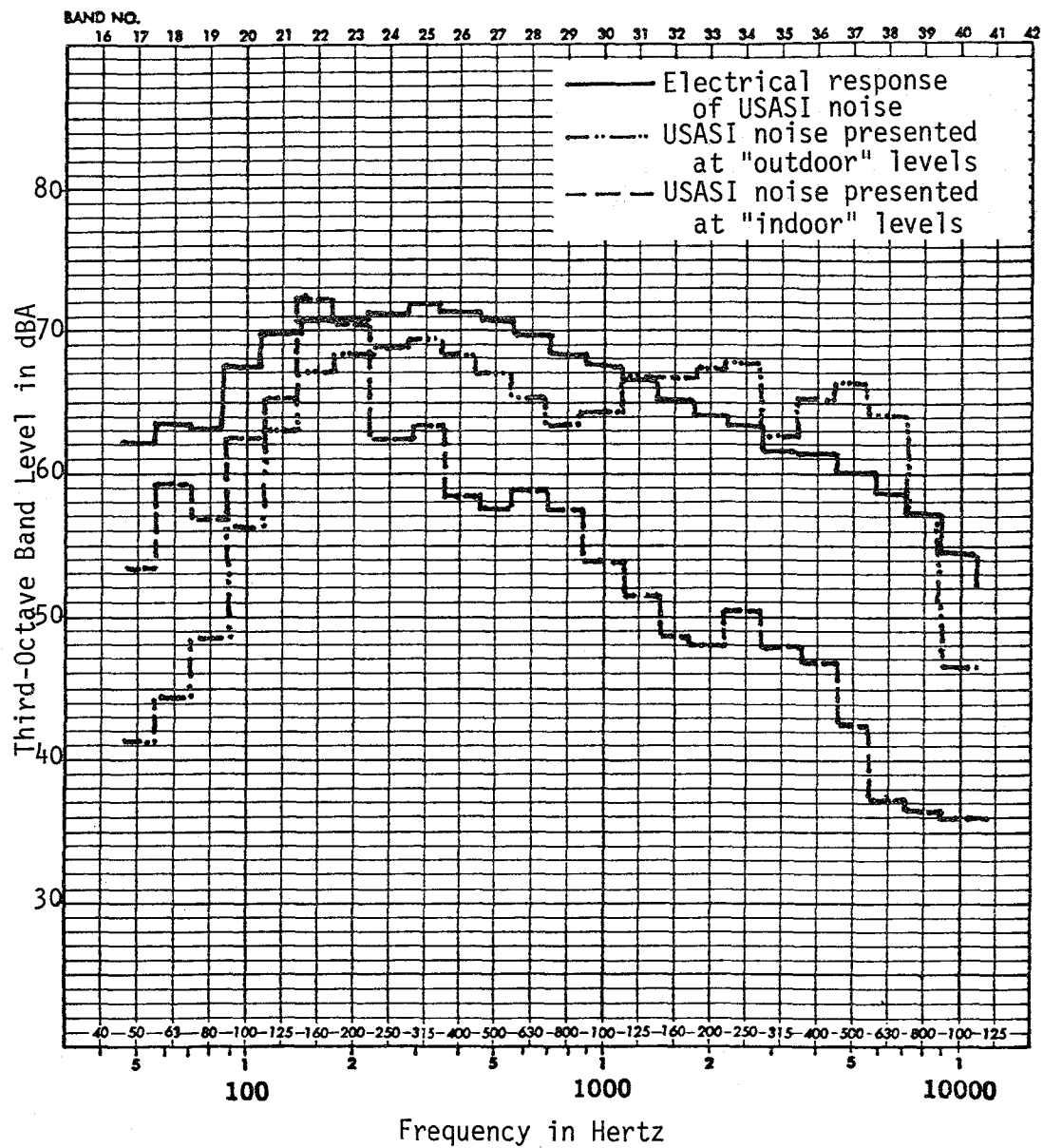


Figure 2-4. Comparison of USASI Noise as Presented to Subjects at the NASA and MAN Facilities.

REFERENCES

- 2-1. Adcock, B.D. and Ollerhead, J.B.: Effective perceived noise level evaluated for STOL and other aircraft sounds. (FAA NO-70-5, May 1970).
- 2-2. Anon: I.S.O. Recommendation R-507: Procedure for describing aircraft noise around an airport. (International Organization for Standardization, June 1970).
- 2-3. Anon: Aerospace Recommended Practice 1071: Definitions and procedures for computing the effective perceived noise level for flyover aircraft noise. (Society of Automotive Engineers, Inc., New York, New York, October 1973).
- 2-4. Anon: Federal Aviation Regulation Part 36: Noise standards: aircraft type and airworthiness certification. (Federal Aviation Administration, March 1978).
- 2-5. Berglund, B.; Berglund, U.; and Lindvall, T.: Scaling loudness, noisiness and annoyance of aircraft noise. (JASA 57/4/930, April 1975).
- 2-6. Boeing Company: Study and development of turbofan nacelle modifications to minimize fan-compressor noise radiation: Vol. VII - Subjective evaluation tests. (NASA CR-1717, January 1971).
- 2-7. Copeland, W.C.T.; Davidson, I.M.; Hargest, T.J.; and Robinson, D.W.: A controlled experiment on the subjective effects of jet engine noise. (Journal of the Royal Aeronautical Society 64/(1)/33, January 1970).
- 2-8. Fletcher, H. and Munson, W.A.: Loudness, its definition, measurement and calculation. (JASA 5/2/82, October 1933).
- 2-9. Goulet, P. and Northwood, T.D.: Subjective rating of broad-band noises containing pure tones. (JASA 53/1/365(A), January 1973).
- 2-10. Hecker, M.H.L. and Kryter, K.D.: Comparisons between subjective ratings of aircraft noise and various objective measures. (FAA NO-68-33, April 1968).
- 2-11. Higgins, T.H.: Human response to sound: the calculation of perceived level, PLdB (noisiness or loudness) directly from physical measures. (FAA RD-76-1, November 1976).
- 2-12. Hinterkeuser, E.G. and Sternfeld, H.: Subjective response to synthesized flight noise signatures of several types of V/STOL aircraft. (NASA CR-1118, August 1968).

- 2-13. ISO/TC 43/WG 12 & 13 (Secretariat 12) 33A: Fourth revised draft of proposed FAA noise certification criteria. (International Organization for Standardization, Technical Committee 43, June 1967).
- 2-14. Kryter, K.D.: Scaling human reactions to sound from aircraft. (JASA 31/11/1415, November 1959).
- 2-15. Kryter, K.D.: Concepts of perceived noisiness, their implementation and application. (JASA 43/2/344, February 1968).
- 2-16. Kryter, K.D.: Possible modifications to the calculation of perceived noisiness. (NASA CR-1636, August 1970).
- 2-17. Kryter, K.D.: A note on the quantity (effective) perceived noisiness and units of perceived noise level. (Journal of Sound and Vibration, 25/3/383, December 1972).
- 2-18. Kryter, K.D.; Johnson, P.J.; and Young, J.R.: Judgment tests of aircraft noise. (NASA SP-189, October 1968).
- 2-19. Kryter, K.D.; Johnson, P.J.; and Young, J.R.: Judgment tests of flyover noise from various aircraft. (NASA CR-1635, August 1970).
- 2-20. Kryter, K.D. and Pearsons, K.S.: Judgment tests of the sound from piston, turbojet and turboprop aircraft. (Sound 1/2/24, March/April 1962).
- 2-21. Kryter, K.D. and Pearsons, K.S.: Some effects of spectral content and duration on perceived noise level. (JASA 35/6/866, NASA TN-D-1873, June 1963).
- 2-22. Kryter, K.D. and Pearsons, K.S.: Modification of noy tables. (JASA 36/2/394, February 1964).
- 2-23. Kryter, K.D. and Pearsons, K.S.: The judged noisiness of a band of random noise containing an audible pure tone. (BBN Report No. 1175, NASr-58, 1965).
- 2-24. Kryter, K.D. and Pearsons, K.S.: Judged noisiness of a band of random noise containing an audible pure tone. (JASA 38/1/106, July 1965).
- 2-25. Kryter, K.D.; Peeler, D.J.; Dobbs, M.E.; and Lukas, J.S.: Reliability of laboratory tests of VSTOL and other loud-duration noises. (NASA CR-2471, November 1974).
- 2-26. Langdon, L.E.; Gabriel, R.F.; and Marsh, A.H.: Investigation of DC-8 nacelle modifications to reduce fan-compressor noise in airport communities: Part VI - psychoacoustic evaluation. (NASA CR-1710, December 1970).

- 2-27. Little, J.W.: Human response to jet engine noises. (Noise Control, 7/3/11, May/June 1961).
- 2-28. Little, J.W. and Mabry, J.E.: Human reaction to aircraft engine noise. (Sound & Vibration, 2/11/14, November 1968).
- 2-29. Little, J.W. and Mabry, J.E.: Empirical comparisons of calculation procedures for estimating annoyance of jet aircraft flyovers. (Journal of Sound and Vibration, 10/1/71, July 1969).
- 2-30. Mabry, J.E. and Little, J.W.: Laboratory study of jet aircraft noise complaint potential as related to context of the noise. (Paper presented at 74th Acoustical Society of America meeting, November 1967).
- 2-31. Mabry, J.E. and Parry, H.J.: An evaluation of psychoacoustic procedures for determining human response to aircraft noise: Vol. II - demonstrated examples. (FAA RD-72-51, II, October 1973).
- 2-32. Mabry, J.E. and Sullivan, B.M.: Response to actual and simulated recordings of conventional takeoff and landing jet aircraft. (NAS1-14404, January 1978).
- 2-33. MAN, Inc.: Noise certification criteria and implementation considerations for V/STOL aircraft. (FAA RD-75-190, November 1975).
- 2-34. MAN, Inc.: Noise certification considerations for helicopters based on laboratory investigations. (FAA RD-76-116, July 1976).
- 2-35. May, D.M. and Watson, E.E.: Correction procedures for aircraft noise data: Volume IV: tone perception. (FAA EE-80-1, Vol. IV, February 1980).
- 2-36. McCurdy, D.A. and Powell, C.A.: Effects of duration and other noise characteristics on the annoyance caused by aircraft flyover noise. (NASA TP-1386, March 1979).
- 2-37. Ollerhead, J.B.: Subjective evaluation of general aviation aircraft noise. (FAA NO-68-35, April 1968).
- 2-38. Ollerhead, J.B.: An evaluation of methods for scaling aircraft noise perception. (NASA CR-1883, October 1971).
- 2-39. Ollerhead, J.B.: Scaling aircraft noise perception. (Journal of Sound and Vibration, 26/3/361, February 1973).
- 2-40. Pearsons, K.S.: The effects of duration and background noise level on perceived noisiness. (FAA ADS-78, April 1966).

- 2-41. Pearsons, K.S.: Assessment of the validity of pure tone corrections to perceived noise level. (NASA SP-189, October 1968).
- 2-42. Pearsons, K.S.: Combination effects of tone and duration parameters on perceived noisiness. (NASA CR-1283, February 1969).
- 2-43. Pearsons, K.S. and Bennett, R.L.: Effects of temporal and spectral combinations on the judged noisiness of aircraft sounds. (FAA NO-69-3, June 1969).
- 2-44. Pearsons, K.S. and Bennett, R.L.: Effects of temporal and spectral combinations on the judged noisiness of aircraft sounds (JASA 49/4/1076, April 1971).
- 2-45. Pearsons, K.S.; Bishop, D.E.; and Horonjeff, R.D.: Judged noisiness of modulated and multiple tones in broad-band noise. (JASA 45/3/742, March 1969).
- 2-46. Pearsons, K.S. and Horonjeff, R.D.: Category scaling judgments tests on motor vehicle and aircraft noise. (FAA DS-67,8, July 1967).
- 2-47. Pearsons, K.S.; Horonjeff, R.D.; and Bishop, D.E.: The noisiness of tones plus noise. (NASA CR-1117, August 1968).
- 2-48. Pinker, R.A.: Mathematical formulation of the Noy tables. (Journal of Sound and Vibration, 8/3/488, November 1968).
- 2-49. Powell, C.A.: A subjective evaluation of synthesized STOL airplane noises. (NASA TN-D-7102, January 1973).
- 2-50. Powell, C.A.: Judgments of relative noisiness of a supersonic transport and several commercial service aircraft. (NASA TN-D-8434, June 1977).
- 2-51. Powell, C.A. and Rice, C.G.: Judgment of aircraft noise in a traffic noise background. (Journal of Sound and Vibration, 38/1/39, January 1975).
- 2-52. SAE Research Project Committee R-6: Development of an aircraft flyover noise rating scale: report on test no. 6: method of adjustment flybys. (Society of Automotive Engineers, Inc., New York, New York, November 1971).
- 2-53. SAE Research Project Committee R-6: Development of an aircraft flyover noise rating scale: report on test no. 7: categorical judgment flybys. (Society of Automotive Engineers, Inc., New York New York, January 1973).
- 2-54. SAE Research Project Committee R-6: Development of an aircraft flyover noise rating scale: report on test no. 9: rating scale improvements. (Society of Automotive Engineers, Inc., New York, New York, December 1973).

- 2-55. Scharf, B. and Hellman, R.: Comparison of various methods for predicting the loudness and acceptability of noise: part II - effects of spectral pattern and tonal components. (EPA 550/9-79-102, November 1979).
- 2-56. Sperry, W.C.: Aircraft noise evaluation. (FAA NO-68-34, September 1968).
- 2-57. Stevens, S.S.: Calculating loudness. (Noise Control 3/5/11, September 1957).
- 2-58. Stevens, S.S.: Perceived level of tones plus noise (Unpublished report for the Laboratory of Psychophysics, Harvard University, 1970).
- 2-59. Wells, R.J.: A new method for computing the annoyance of steady-state noise versus perceived noise level and other subjective measures. (Paper presented at 77th Acoustical Society of America meeting, April 1969).
- 2-60. Wells, R.J.: Calculation of annoyance level of sounds containing multiple pure tones. (Paper presented at the 78th Acoustical Society of America meeting, November 1969).
- 2-61. Wells, R.J.: Jury ratings of complex aircraft noise spectra versus calculated ratings. (Paper presented at 80th Acoustical Society of America meeting, November 1970).
- 2-62. Zwicker, E.: Ein Verfahren zur Berechnung der Lautstärke. (A means for calculating loudness). (Acustica 10/Heft 1/304, 1960).

3.0 EXPERIMENT DESCRIPTION

3.1 SIGNALS

Recordings of commercial jet takeoff and landing manoeuvres were selected from a library of high-quality recordings. Six were chosen which, in the original recording, had a high tone-correction, using a FAR 36 EPNL calculation, six with a low tone correction, and seven with medium correction. These 19 recordings were supplemented with three recordings of the A-300 Airbus, two take-offs and a landing. Ten of the 22 recordings were used twice in the experimental design, to provide a repeatability check (See Table 3-I). These 32 sounds were each presented at five levels, making a total of 160 signals to be judged by each subject. The 160 signals were arranged pseudo-randomly in ten groups of 16, so that no recording of a particular aircraft was presented twice within any group.

3.2 EXPERIMENT DESIGN

The experiment was completed in two main parts, once at MAN in Seattle, where the signals were presented through headphones at levels that could be obtained near an airport in the outdoors, and again at the NASA Langley Research Center, using the interior effects facility where the same signals were presented through loudspeakers situated outside a room of typical construction. Subjects inside the room would thus hear the signals modified in a manner similar to the modification that would occur for real aircraft noises heard from within a typical house situated near an airport.

For both experimental situations, the same design was used. Forty subjects listened to all 160 signals. After reading the instructions and listening to a training tape, each subject heard the standard sound which was given a rating of 10, followed by 16 of the experimental signals, and was asked to compare each signal with the standard. They were required to write down a comparative rating for each signal and also to answer a question on whether they would accept the signal if it were heard four or five times an hour during their waking hours in their homes.

The ten groups of 16 sounds, each group being preceded by the standard, were presented to each subject in one three-hour session. The order of the ten groups was varied to conform with a balanced square design, which gave ten presentation orders. At MAN, subjects were tested one at a time, so each order was used for four subjects. At NASA, four subjects were tested simultaneously, so each order was used once.

TABLE 3-I. EXPERIMENTAL SIGNALS

Operation/Aircraft			Experiment Sound Numbers
1.	Landing	DC-9	1, 11
2.	Landing	DC-9	21
3.	Takeoff	727	22
4.	Takeoff	727	2, 12
5.	Takeoff	DC-8	3, 13
6.	Takeoff	727	23
7.	Landing	727	24
8.	Landing	707	4, 14
9.	Takeoff	707	25
10.	Takeoff	DC-8	5, 15
11.	Takeoff	720	6, 16
12.	Takeoff	DC-8	26
13.	Landing	DC-8	27
14.	Landing	707	7, 17
15.	Landing	DC-8	28
16.	Takeoff	DC-8	8, 18
17.	Takeoff	747	29
18.	Takeoff	707	30
19.	Takeoff	747	9, 19
20.	Takeoff	A-300	31
21.	Takeoff	A-300	32
22.	Landing	A-300	10, 20

3.3 FACILITIES

At MAN, subjects were tested one at a time, in a sound-reduction booth. The recordings were played on a TEAC 3300 tape-recorder into a DBX 122 noise reduction system, a Kenwood KA 5002 amplifier and a set of Koss Pro 4AA headphones, both headphones being provided with the same monaural signal.

At NASA, four subjects were tested in a group, seated in the Interior Effects Room (I.E.R.), a room furnished like a typical living room. The construction of this room is typical of modern single-family dwellings. Four loudspeakers are situated above the ceiling of the room to provide a realistic simulation of aircraft noise in a residential environment. A 4-channel Ampex ATR100 tape recorder was used together with a DBX 154 noise reduction system. The tape recorder was controlled by a PDP-11 computer which also controlled the attenuators that controlled the playback level of the recordings.

3.4 PRESENTATION LEVELS

At both facilities, the signals as presented to the subjects, were aligned on peak dBA, so that at the maximum presentation level the signals would all peak at the same value. At NASA, the "indoor" signals were set to a maximum level of 80 dBA (achieved values ranged from 80.1 to 81.2 dBA), with the lower levels being at 74, 68, 62 and 54 dBA. The standard was presented at 68 dBA. Signal levels were measured using a 1/2" B&K type 4133 microphone situated at head level in the center of the listening facilities, B&K 2606 preamplifier, GenRad 1921 1/3-octave analysis system and PDP-11 computer.

At MAN, using "outdoor" levels, some difficulty was experienced with measuring the levels as presented over the earphones (KOS Pro4AA). Final measurements were made with a GenRad 1560-P83 earphone coupler fitted with a flat plate to adapt it for measurements with circumaural headphones. (See Ref. 3-1) A one inch GenRad microphone, with a GenRad 1982 sound level meter acting as preamplifier, were connected to a GenRad 1921 1/3-octave band multifilter/analyzer and a PDP-11 computer. These measurements gave

an average peak dBA level of 96.5 for the maximum presentation level, with the other levels averaging 90.9, 85.0, 79.2 and 73.3 dBA. The standard was presented at 85 dBA.

3.5 INSTRUCTIONS

The instructions that were given to the subjects were:

INSTRUCTIONS

We are asking you to help answer the question, "How annoying are various kinds of sounds?" we will ask you to listen to some sounds and to rate them in terms of annoyance. The sounds you are to rate will be presented to you one-at-a-time. Listen to all of each sound before making your judgment. In a moment, we will have you listen to a sound with an annoyance score of 10. Use that sound as a standard, and judge each succeeding sound in relation to that standard. For example, if a sound seems twice as annoying as the standard, you will write "20" in the space for that sound on the answer sheet. If it seems only one-quarter as annoying, write " $2\frac{1}{2}$ ". If it seems three times as annoying, write "30". If slightly more than twice as annoying, you may choose to write "21" or "22" or "23", whatever is appropriate. If slightly less annoying than the standard, use the number that best expresses the difference, such as "7" or "8" and so on.

We will also ask you to judge if each sound you hear would be acceptable to you if you experienced it in your home four or five times an hour during your waking hours. This requires a simple "yes" or "no" answer in the space provided on the answer sheet.

Your ratings should reflect only your own opinion of the sounds; that is what we want. Each sound is numbered to correspond to the numbers on your answer sheet.

You will now hear the standard sound with an annoyance rating of 10, followed by five more sounds. Rate each of the sounds following the standard as previously instructed; as score of "20" if twice as annoying, "5" if half as annoying, and so on. Be sure to listen to all of each sound before making your judgment. Also indicate your judgment of the

acceptability of each sound.

* * * * *

3.6 SUBJECTS

Forty subjects were used in both tests. At MAN, half the subjects were female. At NASA, ten of the subjects were male. All subjects were tested audiologically for normal hearing before the test.

3.7 TRAINING

All subjects were asked to read the instructions initially. Then they heard the instructions being read to them on tape, followed by a practice test of the standard sound and five test signals. Their results were then checked by the experimenter for any obvious mistakes; it was found that a small number of subjects would get their answers to the "acceptability" question confused. After the practice test, the ten experimental sessions were administered.

REFERENCES

- 3-1. Michael, P.L. & Bienvenue, G.R.: Calibration data for a circumaural headset (JASA 60/4/944, Oct. 1976).

4.0 DATA ANALYSIS

4.1 NOISE METRIC DESCRIPTION

All the signals used in this experiment were analyzed into 1/3-octave x 1/2-second values, and the resulting time-histories were used to compute metrics for each signal. Unfortunately the signal to noise ratios of the lower level signals acquired at the NASA facility were too low for them to be usable in computation, so the data from the higher level signals were shifted by the difference in presentation level, measured using peak dBA, and the shifted time-histories used in computing the units.

The units are tabulated in Table 4-I. Basically, two weighting procedures were used, PNdB (Ref. 4-1) and dBA (Ref. 4-2). Two methods of detecting tones in a 1/3-octave band spectrum were used, that defined in FAR Part 36 (Ref. 4-1) (which is also that used in ISO R 507), and that in ARP 1071 (Ref. 4-3). Four methods were used to calculate the correction to be added to the weighted sound pressure level for the detected tones. These were that defined in FAR-36 that in R 507 (Ref. 4-4), that in ARP 1071 and that described by J. Little et al (Ref. 4-5). Each correction procedure was applied to each detection procedure. Each of the two weighting procedures was used uncorrected and with each of the correction methods.

Amendment 9 of FAR-36 requires a 5-sample averaging procedure, to account for possible sharing of a tone between adjacent 1/3-octave bands. With aircraft noise, the Doppler shift would ensure that such band-sharing would only happen over a limited time, so the tone-correction is averaged over five successive 1/2-second intervals around the maximum [See FAR-36, Appendix B, §B36.5(n)]. If this averaged value is greater than the value at the peak, the averaged tone correction is used in computing the maximum and the integrated values of tone-corrected PNdB. This calculation was done for PNdB, with the FAR 36 tone detection and correction procedures.

In addition the Kryter and Pearsons tone-correction procedure was used (Ref. 4-6). This varies from the other methods in that it does not compute a correction to be added to each 1/2-second PNL value, but rather corrects the 1/3-octave band level before PNL is calculated.

TABLE 4-I. NOISE METRICS

UNIT #	WEIGHTING	TONE DETECTION PROCEDURE	TONE CORRECTION PROCEDURE	DURATION	NOTES
1	PNL	none	none	m	
2	PNL	none	none	i	
3	PNL	F	F	m	
4	PNL	F	F	i	
5	PNL	F	F	m	5
6	PNL	F	F	i	5
7	PNL	F	R	m	
8	PNL	F	R	i	
9	PNL	F	L	m	
10	PNL	F	L	i	
11	dBA	none	none	m	
12	dBA	none	none	i	
13	dBA	F	F	m	
14	dBA	F	F	i	
15	dBA	F	R	m	
16	dBA	F	R	i	
17	dBA	F	L	m	
18	dBA	F	L	i	
19	PNL	A	F	m	
20	PNL	A	F	i	
21	PNL	A	R	m	
22	PNL	A	R	i	
23	PNL	A	L	m	
24	PNL	A	L	i	
25	dBA	A	F	m	
26	dBA	A	F	i	
27	dBA	A	R	m	
28	dBA	A	R	i	
29	dBA	A	L	m	
30	dBA	A	L	i	
31	PNL	K	K	m	
32	PNL	K	K	i	
33	PNL	F	F	m	H*
34	PNL	F	F	i	H*
35	PNL	K	K	m	N,H
36	PNL	K	K	i	N,H
37	PNL	F	A	m	N
38	PNL	F	A	i	N
39	PNL	F	A	m	N,H
40	PNL	F	A	i	N,H
41	PNL	A	A	m	N
42	PNL	A	A	i	N
43	PNL	A	A	m	N,H
44	PNL	A	A	i	N,H

* * * * *

EXPLANATION OF SYMBOLS USED IN TABLE 4-I.

Tone Detection Procedures: F = FAR-36 (Ref. 4-1)
 A = ARP 1071 (Ref. 4-3)
 K = Kryter & Pearsons (Ref. 4-6)

Tone Correction Procedures: F = FAR 36
 A = ARP 1071
 K = Kryter & Pearsons
 R = R 507 (Ref. 4-4)
 L = Little et al (Ref. 4-5)

Duration: m = maximum 1/2-second value
 i = integrated value

Notes: N - these metric procedures were applied to NASA
 signals only. All others were applied to both.
 5 - these metric procedures used the 5-sample
 averaging technique to account for band sharing.
 H - these metric procedures used the high-frequency
 "cut-off" technique, using only corrections
 applied to bands centered at 1 KHz and above.
 H*- these metric procedures used the high-frequency
 "cut-off" technique for the NASA signals, and an
 approximation thereto for the MAN signals.

* * * * *

Each of the correction procedures was used to compute the maximum 1/2-second value and the value integrated over the range within 10 dB of the maximum value, as specified in FAR 36.

The correction for "pseudo-tones" [FAR 36, Appendix B, §B36.5(m)] was not included in these analyses, as it has been devised to remove any effects of interference between ground reflections and direct sound. These effects occur in monophonic recordings while being much less apparent in real life (stereophonic listening). As the recordings used in the tests were recorded monophonically, the "comb filter" effects of ground reflection interference were present in the play-back sound and were clearly audible. Thus, though the tones are "pseudo" when comparing recordings with "reality" and should be rejected when calculating EPNL for "reality" they were real in this test situation and were therefore not excluded from the calculations.

During the acoustical analyses of the signals presented at NASA it was found that all spectra included a peak in the region of 160-200 Hz, due to the transmission characteristics of the interior effects room. This resulted in many low frequency tone corrections. Some investigators have found over-corrections in the low frequency bands, tones being detected that were not audibly present. It was decided to use the NASA spectra to investigate whether excluding low frequency corrections would have any effect. An arbitrary "cut-off" of 1 KHz was used; corrections in bands centered at 1 KHz or above were included, whereas corrections below these bands were excluded and replaced by any smaller high frequency corrections that the calculation procedures might identify.

This "cut-off" procedure was applied to some of the correction procedures (see Table 4-I), and the statistical analyses showed an improvement in the relationship between the resulting noise metrics and the subjective results. Though the 1/3-octave band data were no longer available for the signals presented at MAN, it was decided to make a crude approximation of the "cut-off" procedure to apply to the MAN data to see if it would produce any effect. This approximation was made by taking the metrics calculated for any flyover, which used a tone correction in a band below

1 KHz, and replacing them by the uncorrected metric (maximum or integrated PNL). These data together with the corrected metrics calculated for the other flyover signals were then used in the statistical analyses.

4.2 PSYCHOPHYSICAL METHOD

The experimental method used was the Magnitude Estimation method. This psychophysical method was introduced by S.S. Stevens (Refs. 4-7 and 4-8) and has been used widely as a method of relating human response evaluations to physical stimuli. Results from a number of studies indicate that the relationship between sensation and the physical stimulus is a power function (Ref. 4-7, p. 166). The relationship is:

$$\psi = kI^n$$

where ψ = subjective response

I = stimulus intensity

k = constant of proportionality

n = constant exponent

If the intensity is expressed in decibels, then the equation after rearranging becomes:

$$\log_{10} \psi = \frac{n}{10} \times \text{dB} + \text{constant}$$

Consequently, a log-log plot of subjective response versus stimulus power gives a linear relation with a slope of $n/10$. The quantity n has been determined experimentally for many stimuli. For noise in particular it has the approximate value of 0.3.

The magnitude estimation method is then utilized to obtain a "Subjective dB" for each noise (Ref. 4-9). Subjective dB is the mechanism for evaluating various engineering calculation procedures. Subjective dB answers the following question: "For a particular engineering calculation procedure as applied to a noise event, do the judges place the noise at the same level as does the engineering procedure and if there is a difference between the judged and calculated level, how great is that difference?" The Subjective dB method for investigating various engineering calculation procedures can

best be understood by reference to Figure 4-1.

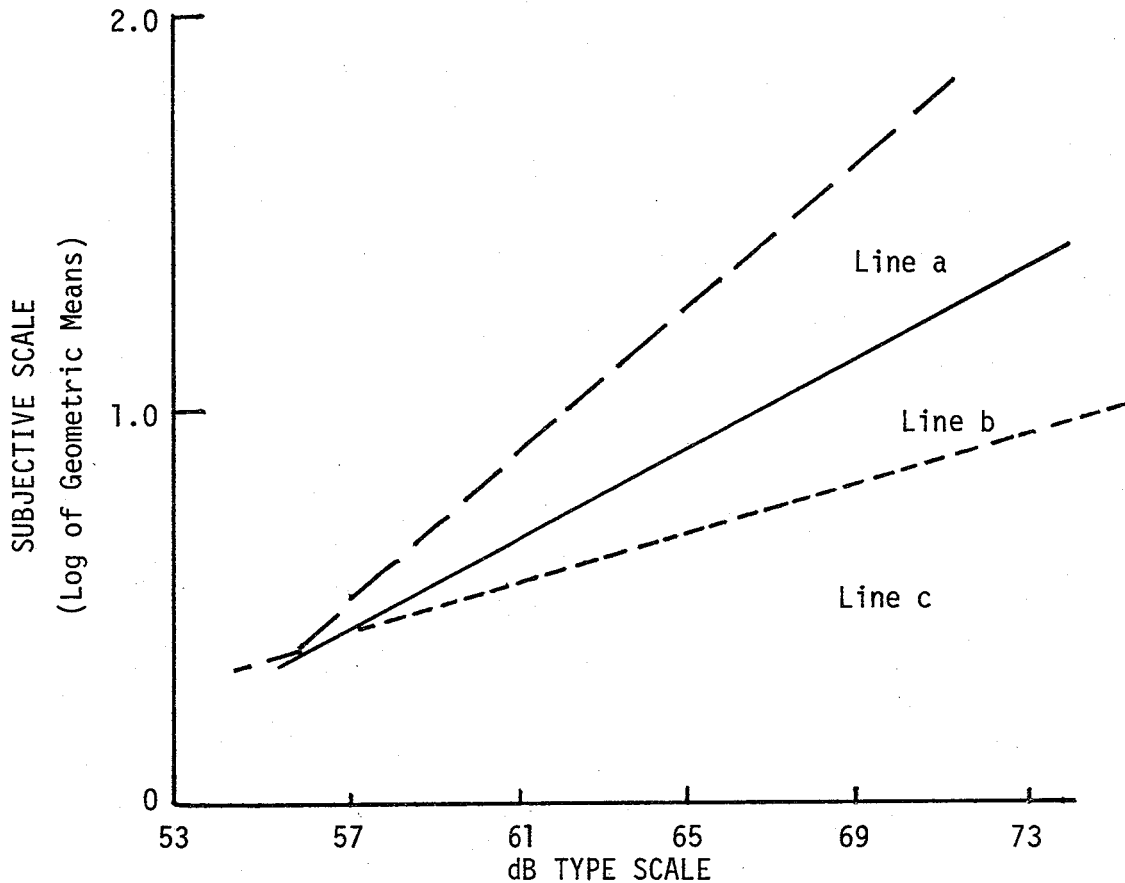


Figure 4-1. Derivation of Subjective dB

Two assumptions form the basis for acquiring a Subjective dB for any one noise. These assumptions are:

That the group of subjects is matching numbers in a manner that reflects the amount of annoyance.

That rate of change of annoyance is different across noises and is a function of a particular noise under investigation.

The abscissa in Figure 4-1 gives values for a particular calculation procedure under investigation while the ordinate represents the evaluations by each judge. Line b is the least squares, best-fitting straight line based on judgments to all noises at all levels. Line b would be based on 108 points, 27 noises at 4 levels. Lines a and c are best-fitting lines

for two hypothetical, individual noises (both Lines a and c would be based on the four levels for a particular noise or on four points).

The operations in calculating a Subjective dB are:

- (1) Obtain equation for best-fitting line using all levels of all noises investigated. This gives an estimate of how well an engineering calculation procedure performs for a wide variety of noises.

- (2) Obtain equation for best-fitting line for each individual noise (Lines a, c, . . .).

- (3) Using the mean of a particular engineering calculation procedure, find, for each individual noise (Lines a and c), the subjective response score predicted by this grand mean.

- (4) Using the subjective response score obtained in (3), calculate the engineering calculation procedure value using best-fitting line based on all observations (Line b). This value is the Subjective dB for ME.

Using results from Figure 4-1 as an example: For the noise on which Line a is based, when the noise is calculated to be at 65 on a dB-type scale, the judge places it at approximately 71, Subjective dB is 71. For the noise on which Line c is based, when the noise is calculated to be at 65 on a dB-type scale, the judge places it at approximately 61, Subjective dB is 61. Each of the 27 noises investigated will be assigned a Subjective dB as described. The predicted results for each engineering calculation system investigated will be similar to results presented in Figure 4-1.

Subjective dB's based on each metric were calculated for each subject for each noise. The resulting tables of 1280 numbers (one table for each metric) were each used in an Analysis of Variance computation to calculate F-ratios for the variance due to the subjects and that due to the noises. An ideal unit would align the noises exactly as the subjects did, and would therefore have an F-ratio not significantly different from 1.0.

REFERENCES

- 4-1. Federal Aviation Administration: Federal Aviation Regulation Part 36 - Noise Standards: Aircraft type and airworthiness certification (March 1978).
- 4-2. Anon: Specification for precision sound level meters. I.E.C. Recommendation Publication 179 (International Electrotechnical Commission, 1965).
- 4-3. Society of Automotive Engineers: Aerospace Recommended Practice 1071 - Definitions and procedures for computing the effective perceived noise levels for aircraft flyover noise. (Oct. 1973) (also ANSI S6.4-1973).
- 4-4. International Organization for Standardization: I.S.O. Recommendation R 507 - Procedure for describing aircraft noise around an airport (June 1970).
- 4-5. ISO/TC 43/WC 12 & 13 (Secretariat 12) 33A: Fourth revised draft of proposed FAA Noise Certification Criteria (June 1967).
- 4-6. Kryter, K.D. and Pearsons, K.S.: Judged noisiness of a band of random noise containing an audible pure tone (JASA 38/1/106, July 1967).
- 4-7. Stevens, S.S.: On the psychophysical law. (Psycho. Rev., 64/3/995, 1957).
- 4-8. Stevens, S.S.: Problems and methods of psychophysics (Psycho. Bull. 55/4/177, 1958).
- 4-9. Mabry, J.E. (for Society of Automotive Engineers): An evaluation of psychoacoustic procedures for determining human response to aircraft noise, Vol. I - Specifications for four experiments. (Report No. FAA-RD-72-51, I, Oct. 1973).

5.0 RESULTS

Using the statistical techniques described in Section 4.1, subjective dB levels for equal annoyance were calculated for each flyover for each of the units studied. These values were entered into an analysis of variance program which computed F-ratios for subjects and noises, and an error term; mean subjective dB levels across 40 subjects were also calculated. For the ideal unit, the F-ratio for noises would be non-significant, as the unit would give the same value to each flyover when judged equally annoying; thus, the range of mean subjective dB's across all 32 flyover signals would be zero.

For 32 noises and 40 subjects, the degrees of freedom for the error are 1209; for $n_1 = 31$ and $n_2 = 1209$, values of the F distribution at the 25, 10, 2.5, and 0.5% points are 1.16, 1.34, 1.57 and 1.79. Thus F-ratios less than 1 are non-significant, and those greater than 2 are highly significant.

For all the units studied, for data collected at both NASA and MAN, the subject F-ratios were non-significant, and the Noise F-ratios were all highly significant (greater than 5). It is therefore apparent that none of the units used was a statistically adequate predictor of subjective annoyance.

Tables 5-I and 5-II give rankings of the noise F-ratios for the data collected at NASA and MAN respectively. The unit numbers are the same as in Table 4-I. Table 5-III compares the relative rank ordering of common units used for both sets of data.

To illustrate how adequately the best units perform in giving equal values to equally annoying levels, Figures 5-1 and 5-2 show the subjective dB levels averaged across 40 subjects for each of the 32 signals, compared with the mean subjective dB level for all signals, for the NASA data (for which the best unit was Unit 44) and for the MAN data (for which the best unit was Unit 32). Signals 1 and 11 are replications of the same flyover, as are 2 and 12, 3 and 13, etc. up to 10 and 20; hence they are grouped in pairs. It is apparent that there is

TABLE 5-I. RANKING OF F-RATIOS FOR DATA COLLECTED AT NASA

Rank	F-Ratio	Unit #	Weighting	Tone Detection Procedure	Tone Correction Procedure	Duration	Notes
1	14.65	44	PNL	A	A	i	N,H
2	15.01	40	PNL	F	A	i	N,H
3	15.12	34	PNL	F	F	i	H*
4	17.73	32	PNL	K	K	i	
5	17.94	10	PNL	F	L	i	
6	18.29	36	PNL	K	K	i	N,H
7	18.60	6	PNL	F	F	i	5
8	18.62	4	PNL	F	F	i	
9	18.64	8	PNL	F	R	i	
10	18.65	38	PNL	F	A	i	N
11	19.10	39	PNL	F	A	m	N,H
12	19.46	33	PNL	F	F	m	H*
13	19.67	24	PNL	A	L	i	
14	20.23	9	PNL	F	L	m	
15	20.50	43	PNL	A	A	m	N,H
16	20.765	20	PNL	A	F	i	
17	20.767	22	PNL	A	R	i	
18	20.80	42	PNL	A	A	i	N
19	21.18	23	PNL	A	L	m	
20	21.54	5	PNL	F	F	m	5
21	21.62	37	PNL	F	A	m	N
22	21.635	7	PNL	F	R	m	
23	21.64	3	PNL	F	F	m	
24	21.78	2	PNL	none	none	i	
25	21.90	41	PNL	A	A	m	N
26 $\frac{1}{2}$	21.92	21	PNL	A	R	m	
		19	PNL	A	F	m	
28	22.43	1	PNL	none	none	m	
29	25.04	17	dBA	F	L	m	
30 $\frac{1}{2}$	26.64	15	dBA	F	R	m	
		13	dBA	F	F	m	
32	27.10	29	dBA	A	L	m	
33 $\frac{1}{2}$	28.20	27	dBA	A	R	m	
		25	dBA	A	F	m	
35	28.88	18	dBA	F	L	i	
36	30.17	14	dBA	F	F	i	
37	30.18	16	dBA	F	R	i	
38	30.32	11	dBA	none	none	m	
39	30.66	30	dBA	A	L	i	
40	32.80	31	PNL	K	K	m	
41	32.927	26	dBA	A	F	i	
42	32.931	28	dBA	A	R	i	
43	33.01	35	PNL	K	K	m	N,H
44	35.16	12	dBA	none	none	i	

TABLE 5-II. RANKING OF F-RATIOS FOR DATA COLLECTED AT MAN

Rank	F-Ratio	Unit #	Weighting	Tone Detection Procedure	Tone Correction Procedure	Duration	Notes
1	5.34	32	PNL	K	K	i	H*
2	6.06	34	PNL	F	F	i	
3	6.71	8	PNL	F	R	i	
4	6.85	4	PNL	F	F	i	
5	6.86	6	PNL	F	F	i	5
6	7.03	22	PNL	A	R	i	
7	7.04	10	PNL	F	L	i	
8	7.05	20	PNL	A	F	i	
9	7.42	24	PNL	A	L	i	
10	11.13	2	PNL	none	none	i	
11	13.55	16	dBA	F	R	i	
12	13.63	26	dBA	A	F	i	
13	13.65	28	dBA	A	R	i	
14	13.72	17	dBA	F	L	m	
15	13.76	14	dBA	F	F	i	
16	13.89	18	dBA	F	L	i	
17	13.91	30	dBA	A	L	i	
18	14.33	1	PNL	none	none	m	
19	14.49	11	dBA	none	none	m	
20	14.66	15	dBA	F	R	m	
21	14.78	13	dBA	F	F	m	5
22	15.00	9	PNL	F	L	m	
23	15.86	5	PNL	F	F	m	
24	15.87	7	PNL	F	R	m	
25	16.01	3	PNL	F	F	m	H*
26	17.47	33	PNL	F	F	m	
27	18.70	12	dBA	none	none	i	
28	21.06	25	dBA	A	F	m	
29	21.07	27	dBA	A	R	m	
30	22.38	29	dBA	A	L	m	
31	22.75	21	PNL	A	R	m	
32	22.76	19	PNL	A	F	m	
33	23.98	23	PNL	A	L	m	
34	32.62	31	PNL	K	K	m	

* * * * *

EXPLANATION OF SYMBOLS USED IN TABLES 5-I and 5-II

Tone Detection Procedures F = FAR-36 (Ref. 4-1)
 A = ARP 1071 (Ref. 4-3)
 K = Kryter & Pearsons (Ref. 4-6)

Tone Correction Procedures: F = FAR-36
 A = ARP 1071
 K = Kryter & Pearsons
 R = R 507 (Ref. 4-4)
 L = Little et al (Ref. 4-5)

Duration: m = maximum 1/2 second value
 i = integrated value

Notes: N - these metric procedures were applied to NASA
 signals only. All others were applied to both.
 5 - these metric procedures used the 5-sample
 averaging technique to account for band sharing.
 H - these metric procedures used the high frequency
 "cut-off" technique, using only corrections
 applied to bands centered at 1 KHz and above.
 H*- these metric procedures used the high frequency
 "cut-off" technique for the NASA signals, and
 an approximation thereto for the MAN signals.

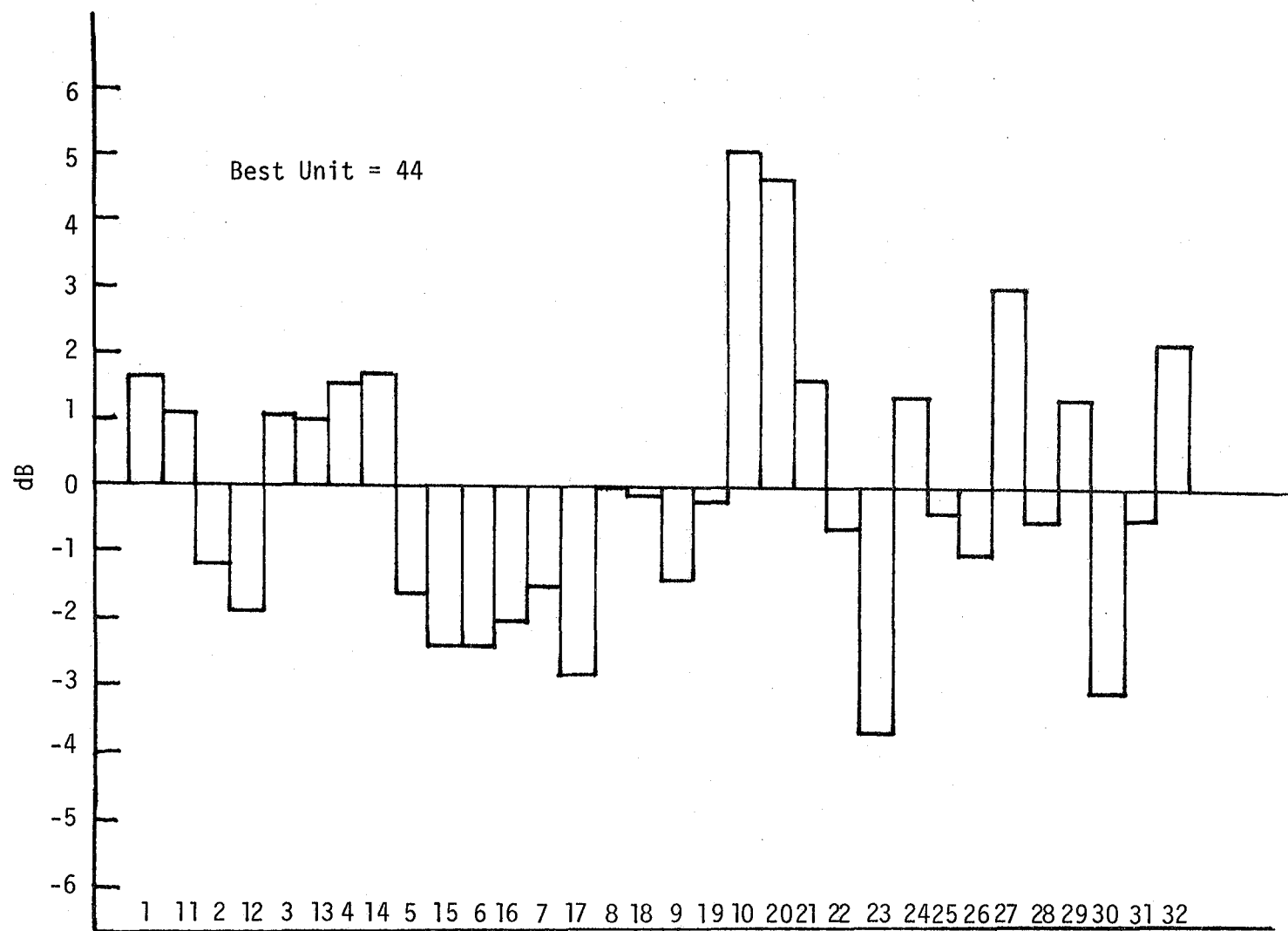
* * * * *

TABLE 5-III. RELATIVE ORDERING OF F-RATIOS
FOR COMMON UNITS (UNIT NUMBERS)

MAN Data	NASA Data	MAN Data	NASA Data
32	34	1	19
34	32	11	1
8	10	15	17
4	6	13	15
6	4	9	13
22	8	5	29
10	33	7	27
20	24	3	25
24	9	33	18
2	20	12	14
16	22	25	16
26	23	27	11
28	7	29	30
17	3	21	31
14	5	19	26
18	2	23	28
30	21	31	12

a wider spread of unit values in Figure 5-1 (NASA data) than in Figure 5-2 (MAN data); this is reflected in the corresponding range of mean subjective dB values (8.68 and 6.43) and the F-ratios (14.65 and 5.34). The F-ratio for the best unit for the MAN data is very much smaller than that for the NASA data, though the range of subjective dB values is only 2.25 dB smaller.

The accuracy of the experimental technique is reflected in the comparisons of the replications; the largest difference in mean subjective



Subjective dB for 32 experimental signals

Figure 5-1. Results from NASA

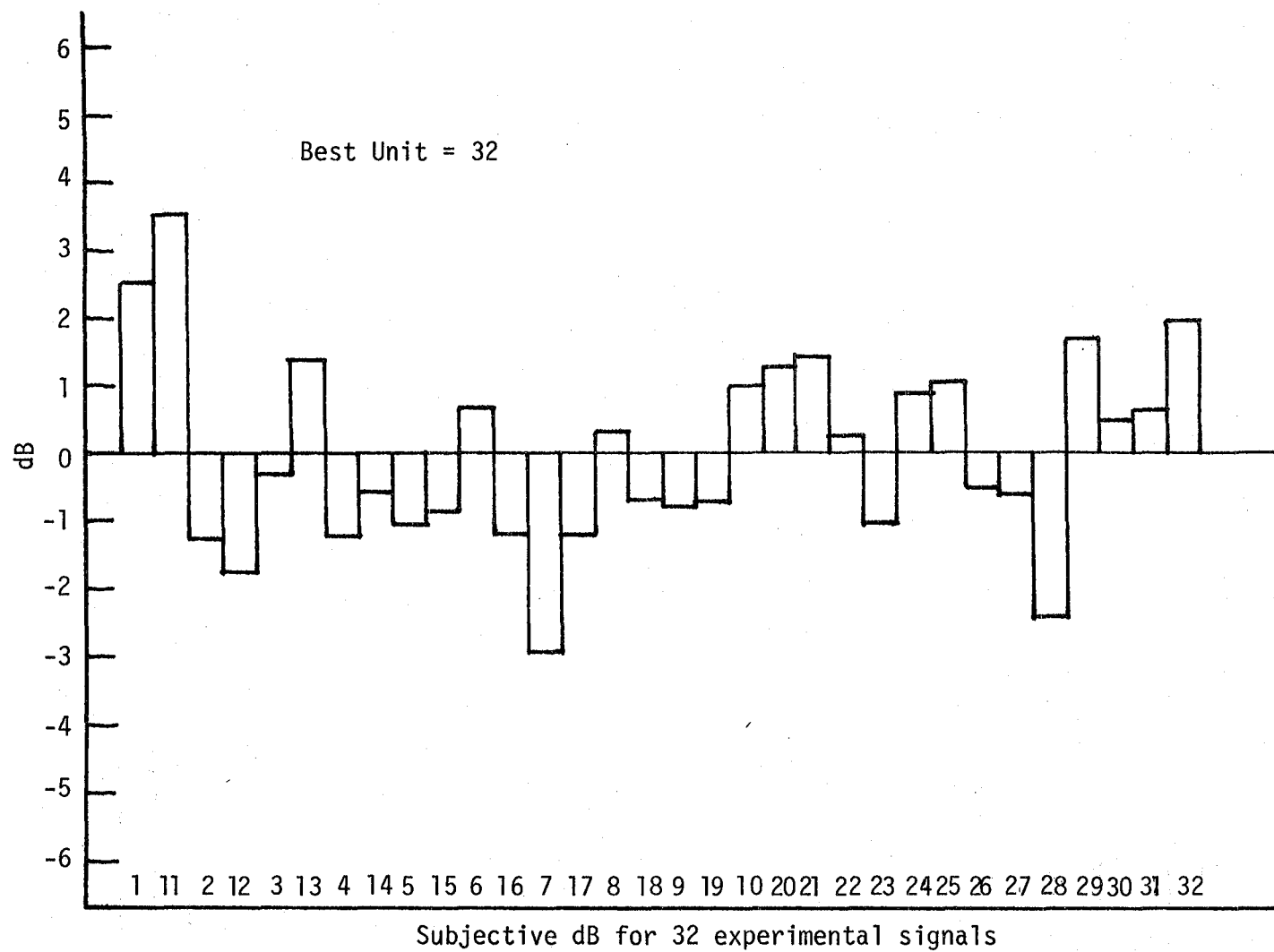


Figure 5-2. Results from MAN

dB for a repeated signal is 1.88 dB, which compares with previous studies using this technique, in which a difference of up to 2 dB has been found to be normal for repeated signals.

Considering the F-ratio ranking tables, some conclusions can be drawn about the calculation procedures used. No one unit stands out as clearly better than the rest, but comparisons do show some definite trends. The main variables in the calculation procedures used are the weighting method (dBA or PNL), the tone detection procedures (FAR-36 [which is also ISO R-507], ARP 1071 and Kryter & Pearsons), the tone correction procedure (FAR-36, ARP 1071, R-507, Little and Kryter & Pearsons), and the duration (maximum 1/2 second or integrated value), with, in addition, extra procedures such as the 5-sample tone correction, or the inclusion of only high frequency tone corrections (See Section 4.1).

To demonstrate whether any of the techniques studied improved the units, Average F-ratios were obtained for units with one technique in common and for the equivalent units with another technique. If one average F-ratio value is less than the other, then in general that technique improves the predictive ability of the units. For example, units using the FAR-36 detection technique are compared with the equivalent units using the ARP 1071 detection technique. All other conditions are held constant except the two detection procedures. A summary of these comparisons of the Average F-ratios for the MAN and NASA results is given in Table 5-IV.

For the data collected at MAN, using "outdoor" signal levels and spectra, Table 5 - II shows the PNL weighting giving better results (smaller F-ratios) than the dBA weighting (other variables being constant) for all units; the average F-ratio for 14 PNL-based units is 13.138 compared to 15.946 for the comparable 14 dBA-based units (See Table 5-V). The integrated units performed substantially better than maximum 1/2 second values (the average F-ratios for 17 comparable units being 10.152 for integrated units and 18.754 for maximum units) (Table 5-IV). Comparing the FAR-36/ISO R-507 detection method with that of ARP 1071, the

TABLE 5-IV. AVERAGE F-RATIOS FOR TWO SETS OF DATA

	Calculation Procedure		Duration Correction		Tone Detection Procedure	
	dBA	PNL	Integrated	Max. $\frac{1}{2}$ Sec.	FAR-36	ARP-1071
MAN results:	15.95	13.14	10.15	18.75	12.65	16.39
NASA results:	29.50	20.65	22.64	24.23	22.54 *(18.47	24.01 22.73)
*Results based on only PNL comparison for NASA data.						
Tone Correction (FAR-36 Detection, Integrated PNLT)						
	R-507	FAR-36	FAR-36 (5-sample average)	J.Little	Kryter/ Pearsons	
MAN results:	6.71	6.85	6.86	7.04	5.34	
NASA results:	18.64	18.62	18.60	17.94	17.73	

FAR-36 version worked better, average F-ratios being 12.653 for FAR-36 and 16.391 for ARP 1071, using 12 comparable units (Table 5-IV).

The standard tone correction methods used for the MAN data show little difference (Table 5-IV). Looking at integrated PNLT with the FAR-36 detection method, the F-ratios are 6.71 for the R-507 correction, 6.85 for the FAR-36 correction, 6.86 for the FAR-36 correction with the 5-sample average correction, and 7.04 for Little's correction. The best unit used with this data was integrated PNL plus the Kryter and Pearsons tone correction, with an F-ratio of 5.34.

For the data collected at NASA using "indoor" levels and spectra, Table 5-I shows PNL performing substantially better in rank than dBA; again averaging across 14 comparable units, the average F-ratio for the

PNL-based units is 20.654, compared with 29.504 for the dBA-based units (Table 5-IV). Comparing duration corrected with maximum values gives a less clear-cut result, the average F-ratios being 22.635 for the 22 integrated units and 24.229 for 22 maximum units, showing a slight improvement with integration. However, using only the 15 PNL-based units gives average F-ratios of 18.469 for integrated units and 22.727 for maximum units, a clearer indication of the efficacy of the duration correction (Table 5-IV).

Comparing the detection methods, the average F-ratios for 16 comparable units are 22.540 for the FAR-36 method and 24.006 for the ARP 1071, a slight edge for the FAR-36 version. These results show the same trend as the results from the MAN data.

The standard tone corrections again showed practically no differences; for integrated PNLT using the FAR-36 detection method, the F-ratios are 18.60 for FAR-36 with the 5-sample correction, 18.62 for FAR-36 (without it), 18.64 for ISO R-507, 18.65 for ARP 1071 and 17.94 for Little's procedure. Again Kryter and Pearsons' procedure performed better than the other well known methods; the F-ratio for the integrated unit corrected this way is 17.73 (Table 5-IV).

From these results, it is evident that the data collected at NASA differs from that collected at MAN, in that the F-ratios are much larger for the NASA data. To investigate this phenomenon further, the units described in Section 4.1, in which tone corrections were only applied to bands of 1 KHz and above, were computed. Again comparing F-ratios averaged across equivalent units, the effect of this "cut-off" was to improve the parameter from 20.137 to 17.447; looking at integrated PNLT with standard procedures (thus excluding Kryter and Pearsons' method), the improvement is from 19.357 to 14.927. The "high frequency only" correction degrades Kryter and Pearsons' method but improves the other procedures remarkably. An approximate version of the "high frequency only" rule applied to integrated PNL plus the FAR-36 procedure for the MAN data altered the F-ratio from 6.85 to 6.06, a small change but again in the direction of improvement.

Considering only the standard procedures for calculating EPNL for the two sets of data, the F-ratios are, for the NASA data, 18.60 for FAR-36, 18.64 for ISO R-507 and 20.80 for ARP 1071, and for the MAN data, 6.86 for FAR-36 and 6.71 for ISO R-507.

Table 5-V shows the range of tone correction values for different detection and correction procedures, calculated from the difference between EPNL and PNLD, for the 32 flyover signals (averaging values for the five presentation levels for each signal) for the two sets of data. The values for the Kryter and Pearsons procedure, which performed well in both cases, have the widest range, but, more significant perhaps, go down to a zero correction. The versions in which high frequency corrections only were applied have the same upper values as their more usual counterparts but also go down to zero. Other correction procedures, such as ISO R-507 and ARP 1071, which give a zero correction for low intensity tones, did not go to zero for these average values because at least one of the presentation levels gave a measurable correction.

TABLE 5-V. TONE CORRECTION RANGES

Detection Procedures	Correction Procedures	NASA data	MAN data
FAR-36	FAR-36	1.0-3.5	1.0-4.5
	FAR-36 (+5-sample correction)	1.0-3.5	1.0-4.5
	ISO R-507	1.0-3.5	0.5-4.5
	ARP 1071	1.0-3.5	0.5-4.5
	Little	1.0-3.5	0.5-4.5
	FAR-36 (high frequency only)	0-3.5	0-4.5
	ARP 1071 (high frequency only)	0-3.5	
ARP 1071	FAR-36	1.5-4.5	1.0-6.5
	ISO R-507	1.5-4.5	1.0-6.5
	ARP 1071	1.5-4.5	
	Little	1.5-4.5	1.0-7.5
	ARP 1071 (high frequency only)	0-4.5	
Kryter & Pearsons	Kryter & Pearsons	0-6.5	0-7.5

It would therefore seem reasonable to suggest that a better tone correction procedure than any at present standardized would reduce the tone penalty for the low frequency bands, though the question as to the adequacy of the detection procedure remains open.

All of the NASA flyover presentations had relatively high tone corrections (1 dB or above), which occurred in many cases in the low frequency bands due to the filtering characteristics of the room. The effect of this filtering is shown in Figures 5-3 to 5-6. In Figure 5-3, the position of the tone has not been affected by the filtering; it remains in the 2.5 KHz band. Neither has it been affected in Figure 5-4; it remains in the 200 Hz band. However, Figures 5-5 and 5-6 demonstrate cases where the tone has moved to low frequency bands.

The difference in the results from the two sets of data studied here may be attributed at least in part to two factors: the lower presentation levels and the relatively greater low frequency energy in the flyovers used at NASA by comparison with those used at MAN. The NASA results show PNL to be superior to dBA more clearly than the MAN results, which may be due to the differences in weighting of the low frequency energy, which would be more apparent with the greater low frequency energy.

The MAN results show the need for the duration correction more readily than the NASA results; this may be due to the higher levels used at MAN.

The tone correction procedures do not differ so markedly, though they do improve the prediction of subjective reaction over the uncorrected versions (more clearly in the NASA data; for integrated PNLT in the MAN data). The present tone corrections would seem to work best for high intensity, high frequency tones.

The effect of the ad-hoc "high frequency only" cut-off improved the integrated PNLT units for both sets of data.

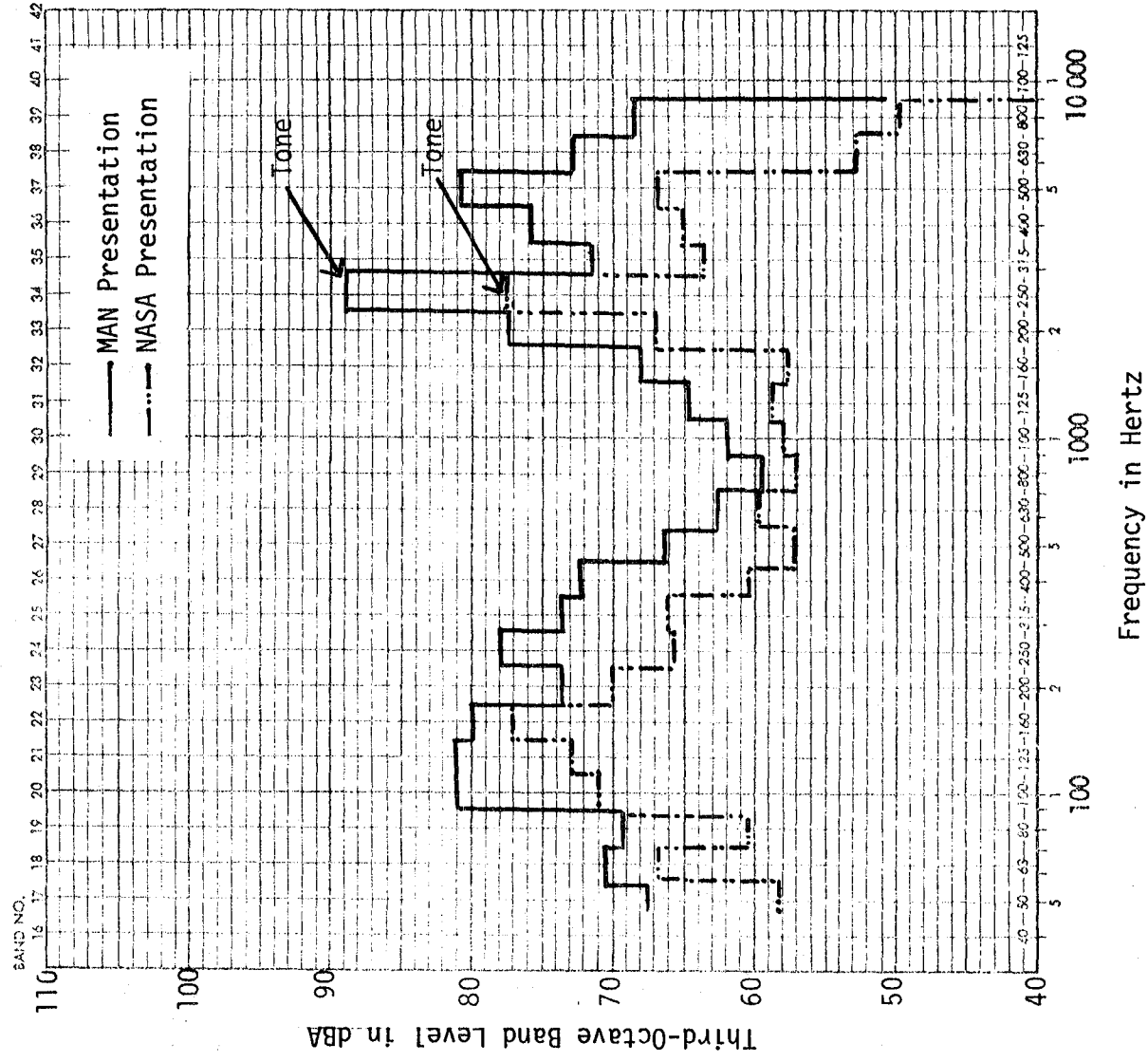


Figure 5-3. Signal 7 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations

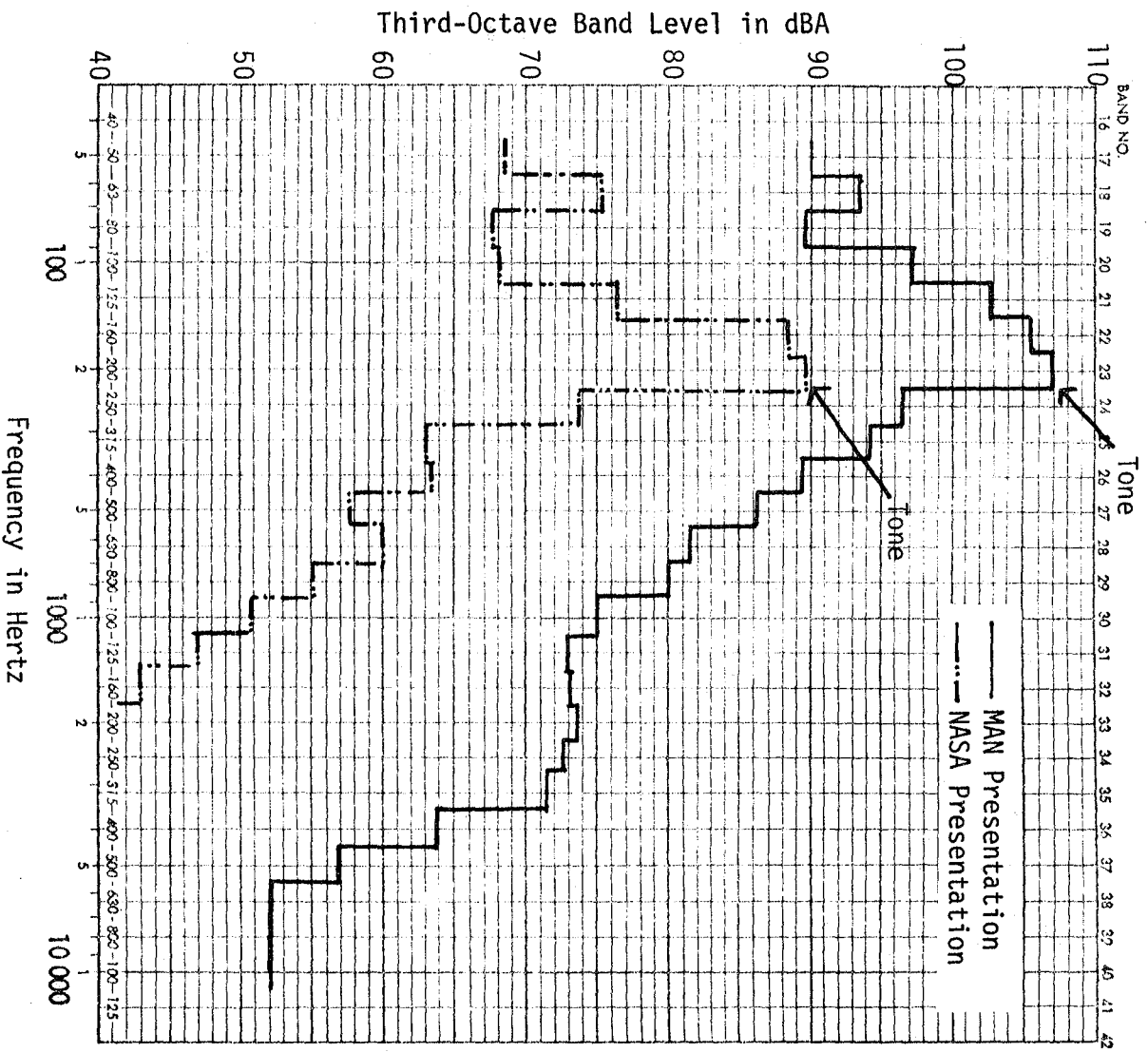


Figure 5-4. Signal 30 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations

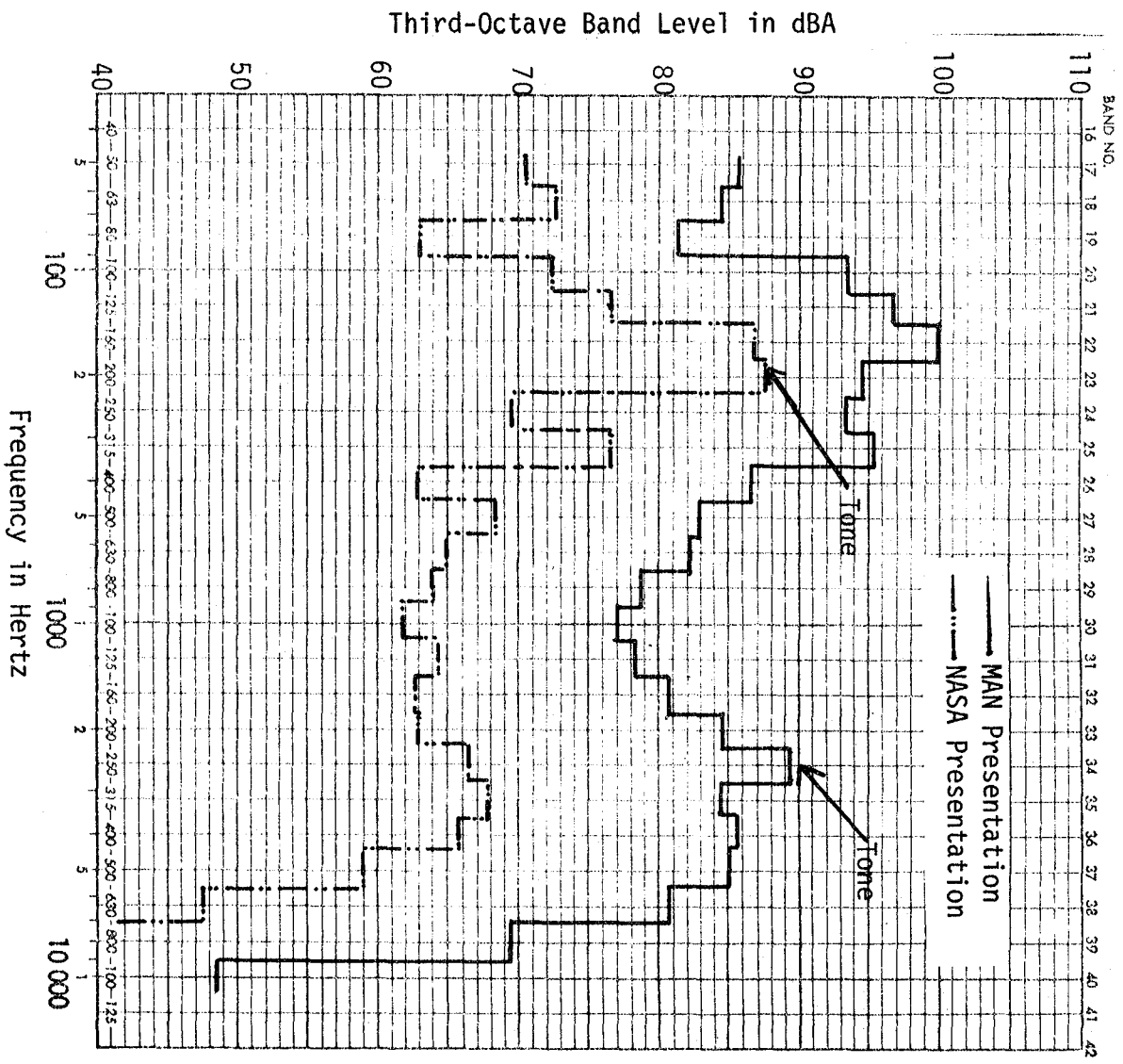


Figure 5-5. Signal 21 at level 1: Comparison of Peak Spectra for MAN and NASA Presentations

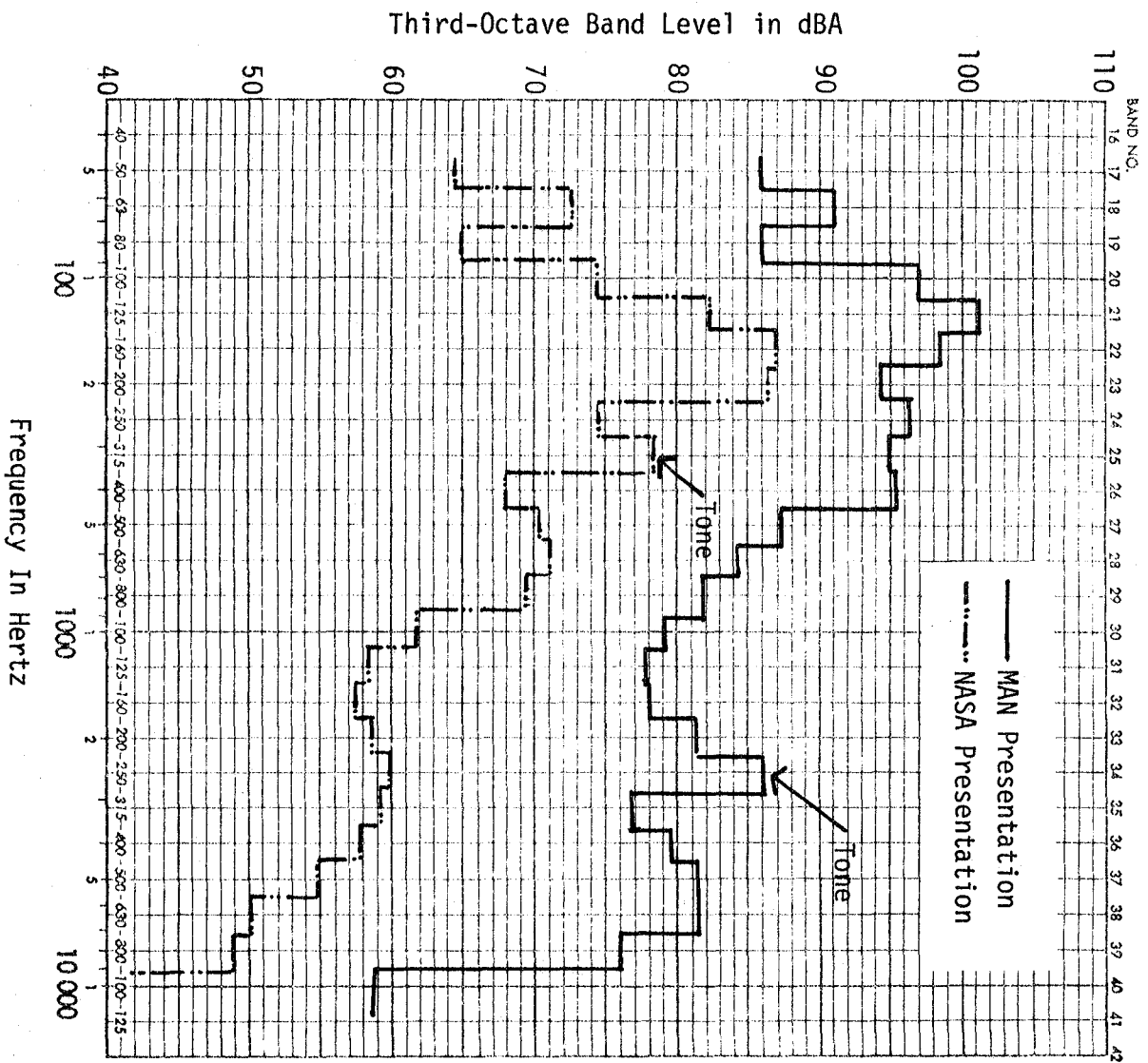


Figure 5-6. Signal 24 at Level 1: Comparison of Peak Spectra for MAN and NASA Presentations

6.0 SUMMARY OF CONCLUSIONS

- Both sets of results show that the PNL calculation procedure has a higher relationship to the judgment data than does dBA and is thus a more valid transformation on the acoustical data.
- The integrated duration correction is more effective than the maximum 0.5 second duration correction approach. This finding was supported by results from both the NASA and MAN studies.
- For the tone detection procedure, the FAR-36 method is superior to the ARP 1071 approach. This finding was not as definite for the NASA results as for those from MAN. However, for comparisons based only on PNL (dBA omitted), the difference in favor of the FAR-36 method was increased for the NASA results.
- For the five tone correction procedures investigated, differences were minimal for both sets of results. However, the Kryter and Pearsons approach was slightly better than the other four procedures, particularly for the MAN results.
- Omitting tone corrections below 1 KHz appreciably increased the relationship between the judgment and acoustical data for the NASA study but to a much smaller extent for the data collected at MAN. This is attributed to the fact that the filtering effect of the room (indoor listening) shifted the identified tone to the low frequency bands to which the listeners did not find annoying.
- The MAN results indicate a requirement for a duration correction to a greater extent than those based on the NASA study. Since considerably higher noise levels (outdoor listening) were investigated at MAN, it is likely that the duration factor is more significant at levels typical of outdoor noise exposure.

•A summarizing conclusion is that the tone corrections presently in use are most effective in the evaluation of high intensity noise containing higher frequency tones. Outdoor aircraft noise is more validly measured relative to human response than is typical indoor aircraft noise.

1. Report No. NASA CR-165910	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle A Study of Noise Metric and Tone Correction Accuracy		5. Report Date April 1982	
		6. Performing Organization Code	
7. Author(s) B.M. Sullivan and J.E. Mabry		8. Performing Organization Report No. MAN 1046	
9. Performing Organization Name and Address MAN-Acoustics & Noise, Inc. 16201 39th NE Avenue Seattle, WA 98155		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. NAS1-15955	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		13. Type of Report and Period Covered Contractor Report Oct. 1979 - Feb. 1982	
		14. Sponsoring Agency Code	
15. Supplementary Notes Langley Technical Monitor: Clemans A. Powell, Jr. Final Report			
16. Abstract As a means of investigating methods currently used to measure human response to aircraft flyover noise, two laboratory type studies were completed. One study was aimed at obtaining response to high level aircraft noise usually experienced in the outdoor environment; this study was completed in the contractor's laboratory in Seattle, WA. A second study investigated response to aircraft flyover noise typical of indoor exposure and was completed in the NASA Langley Research Human Effects Laboratory. Flyover signals and methods were identical for the two studies with the exception that the flyover signals were filtered by a typical house construction at NASA. It was concluded that current methods for evaluating response to aircraft flyover noise are more accurate for outdoor noise (high intensity noise containing higher frequency tones) than for indoor noise.			
17. Key Words Psychoacoustics Aircraft noise Noise calculation procedures		18. Distribution Statement Unclassified - Unlimited Subject Category - 45	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 72	22. Price

End of Document